Detection of somatic mutations in *EGFR* and *KRAS* genes in formalin-fixed paraffin-embedded specimens from Non-Small Cell Lung Cancer patients

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science in Biological Science

at

The University of Waikato

by

TANIA ALICE LAW

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2012

Abstract

Non-small-cell lung cancer (NSCLC) accounts for 80% of lung cancer and has a 5 year survival rate of just 15%. Recently, tyrosine kinase inhibitors (TKIs) have been found to give dramatic therapeutic benefit to some patients increasing both their overall survival and quality of life compared with the more traditional and aggressive platinum-based chemotherapy. Subsequently, it was shown that *EGFR* and *KRAS* mutations act as biomarkers in predicting of patient response or lack thereof, respectively to TKIs.

As such, the establishment of a clinically relevant assay for the detection of *EGFR* and *KRAS* mutations would contribute to personalised treatment of NSCLC patients to ultimately improve patient outcomes. However, there are many limitations and obstacles that limit the sensitivity of current protocols to detect these biomarkers, including the inherent heterogeneous nature of NSCLC tumours and the poor quality of genomic DNA due to biopsies being stored formalin-fixed, paraffin-embedded (FFPE) specimens.

This study investigated the use of mineral oil to produce a high yield of quality DNA from FFPE NSCLC specimens. Additionally, significant optimisation experiments were carried out to find the optimum conditions for efficient amplification during PCR thermocycling using DNA derived from these FFPE specimens. Subsequently, the use of standard PCR, Co-amplification at lower denaturation temperature (COLD) PCR, restriction digest-mediated mutant enrichment PCR, single-stranded conformation polymorphism (SSCP) and DNA melt curve analysis were investigated as *EGFR* and *KRAS* mutation detection protocols in NSCLC FFPE specimens.

Significantly, this study found that the use of mineral oil contributed to the extraction of a high yield of quality genomic DNA from the FFPE specimens and the use of a high fidelity DNA polymerase enzyme and a PCR buffer with a high magnesium concentration were required to produce amplifiable products from the FFPE specimens. Subsequently, it was found that both SSCP and DNA melt curve analysis could detect putative mutations in *EGFR* exon 21 and 19.

Given that SSCP and DNA melt curve analysis could detect mutations in the NSCLC FFPE specimens, these mutation detection protocols were shown to be more sensitive than direct sequencing. However, further work is required to establish a clinical relevant mutation detection protocol for the routine detection of *EGFR* and *KRAS* mutations in FFPE NSCLC specimens.

Acknowledgements

Firstly, I must acknowledge the guidance, encouragement and expertise of my supervisor, Dr Ray Cursons. It has been an amazing experience working under your supervision which has seen me grow both as a researcher and a person. Thank you for your patience, your understanding and most of all for challenging me and continually leading by example.

I must also acknowledge the exceptional group of people whom have shared this journey with me, Rebecca, Nadine, Claire and Jo. I couldn't have wished for a greater bunch of people to have gone through the craziness that has been our Masters. The friendship and memories (what happens in Queenstown stays in Queenstown) we have shared have made this endeavour such an unique experience. To Olivia and Greg, thank you for all your help and patience, and the constant smiles in the lab.

Also, a big thank you to Greg, Justin and the rest of the Sir Edmund Hillary Scholarship support network. The financial support I have received during my studies has been superseded by the personal support you have shown me during both my undergrad and postgrad studies.

And finally, to my family and friends, thank you for the constant encouragement and support, both in the good times and the bad. I couldn't have done this without you all.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

NSCLC	Non-small cell Lung Cancer
SCLC	Small-cell Lung Cancer
EGFR	Epidermal Growth Factor Receptor
KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
TKIs	Tyrosine Kinase Inhibitors
PCR	Polymerase Chain Reaction
FFPE	Formalin-fixed and Paraffin-embedded
DNA	Deoxyribonucleic acid
bp	Base pair
RCF	Relative Centrifugal Force
ATP	Adenosine triphosphate
OS	Overall survival
RR	Response Rate
PFS	Progression Free Survival
TK	Tyrosine kinase
ssDNA	single-stranded DNA
NZ	New Zealand
GTP	guanosine triphosphate
GDP	guanosine diphosphate
HRM	High resolution DNA melting
SSCP	Single-stranded Conformation Polymorphism
PAGE	Polyacrylamide gel electrophoresis
COLD-PCR	Com-amplification at lower denaturation temperature PCR

FDA	Food and Drug Administration
NNK	nicotine-derived nitroaminoketone
GAPs	GTPase-activating proteins
kDa	kilo Daltons

Chapter 1 General Introduction & Literature Review

1.1 General Introduction

Lung cancer is the cause of the highest rate of cancer-associated mortality in the world, with 1.38 million deaths worldwide annually (Hirsch, 2009; Lee et al., 2011). Significantly, it is ranked second, behind only cardiovascular disease, as the leading cause of death in New Zealand (NZ) and has accounted for more than a quarter of all deaths since the 1990s (Hodgen, Tobias, & Chenung, 2002). Compared with European countries and Australia, NZ has a relatively high lung cancer incidence and mortality rate, with a growing epidemic among females and Maori, as well as lower socioeconomic groups (Blakely, Shaw, Atkinson, Cunningham, & Sarfati, 2011; Stevens, Stevens, Kolbe, & Cox, 2008; Sutherland & Aitken, 2008). Moreover, while the NZ government has forecasted that overall rates of lung cancer will decline over the next several years, the cost associated with its prevention and treatment will remain a substantial burden to the health system for the foreseeable future (Masago et al., 2008).

Traditionally, lung cancer is treated with combination platinum-based chemotherapy, yet this offers only a modest increase in overall survival while adversely affecting the patients' quality of life (Triano, Deshpande, & Gettinger, 2010). Accordingly, the response rate to standard platinum-based chemotherapy is typically 20-30% with a median survival time of approximately 4 months (Mitsudomi, 2010; Pao et al., 2009; Sriram, Larsen, Yang, Bowman, & Fong, 2011). This minimal disease response shown by 70-80% of patients can partly be attributed to the aggressive nature of the disease, and the fact that most patients present with an advanced metastatic stage at diagnosis. However, the mechanisms that contribute to the poor treatment response are wide-ranging and multi-factorial.

Consequently, over the past several years, researchers have suggested many genetic, environmental, hormonal and viral factors as risk factors for lung cancer,

particularly amongst non-smokers. Whilst there have been many associations between risk factors and lung cancer incidence, no dominant risk factor has been validated across racial and ethnic groups (Abidoye, Ferguson, & Salgia, 2007; Lee, et al., 2011). Indeed, despite tobacco smoking being the predominant risk factor, only 10-20% of smokers will develop lung cancer, highlighting the multifaceted etiology of the disease (Sriram, et al., 2011; Yano, Haro, Shikada, Maruyama, & Maehara, 2011).

In fact, this is typical of complex disease phenotypes, whereby a single causal factor cannot explain the multi-step and complicated process of tumourigenesis and lung cancer development. Instead, complex gene-environment interactions are implicated in the carcinogenesis of lung cancer, and with the completion of the Human Genome Project, there is a better understanding of the molecular origins and evolution of the disease (Herbst, Heymach, & Lippman, 2008).

1.2 Biology of Lung Cancer

Lung cancer, like most cancers, is a conglomeration of diseases, broadly divided into two main groups, small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC), based on the different tissues of origin from which the disease develops (Johnson & Kelley, 1995). SCLC which makes up 10-15% of lung cancers is of neural crest origin and generally initially responses well to chemotherapy, although reoccurrences are common (C. H. Lin, Yeh, Chang, Hsu, & Chang, 2010; Sharma, Bell, & Settleman, 2007). Whereas, NSCLC makes up 85-90% of reported lung cancers and while it originates from lung epithelial cells, there are diverse histological subtypes including adenocarcinoma, bronchioalveolar, squamous, anaplastic and large-cell carcinomas (Cataldo, Gibbons, Perez-Soler, & Quintas-Cardama, 2011).

Given that survival with lung cancer is closely associated with tumour stage and treatment, therapeutic treatment and prognosis has traditionally depended on the histological subtype and stage of cancer, as well as the patient's general wellbeing (Stevens, et al., 2008). Significantly, the most frequent type of histological tumour is NSCLC adenocarcinoma, which account for 40% of all cases of lung cancer in North America and is the predominant form of the disease in most of the world

(Jemal, Siegel, Xu, & Ward, 2010). This subtype has a better prognosis than other lung cancers, and with optimal treatment the possibility of a cure is greatest for this group of patients (Stevens, et al., 2008).

As such, significant research has been undertaken to establish treatment regimens that offer optimal therapeutic benefit for lung cancer patients. Accordingly, recent studies have focused on the unique biology of lung cancer tumourigenesis, in attempt to establish the pathogenesis pathways that lead to the development of the malignant phenotype (Igbokwe & Lopez-Terrada, 2011). In particular, focus has been on understanding the complex and aberrant activation of signalling pathways that stimulate that six hallmarks of cancers cells; uncontrolled proliferation, evasion of apoptosis, insensitivity to anti-growth signals, sustained angiogenesis, tissue invasion and metastasis, self-sufficiency in growth signals (Gazdar, 2010).

1.3 Non-small-cell Lung Cancer

As NSCLC makes up the majority of lung cancers, it contributes significantly to the high mortality rate associated with the disease. NSCLC tumours do not respond well to chemotherapy, partly due to the advanced stage at which most patients present at diagnosis (John, Liu, & Tsao, 2009). As such, the median survival associated with the disease is only 4 to 5 months and a prognosis of metastatic NSCLC remains poor, with a 5-year survival of only 5% (Cataldo, et al., 2011).

Indeed, NSCLC is a highly aggressive disease and is present both in smokers and non-smokers (Wakelee et al., 2007). NSCLC is relatively asymptomatic in that patients often only notice a persistent cough, which is not attributed to suffering from the disease. Further, while cigarette smoking remains the most substantial contributor to the development of lung cancer, 15% of males and 53% of females, which combined accounts for 25% of all lung cancers worldwide, diagnosed with the disease are non-smokers, highlight the multi-faceted etiology of the disease (Sun, Schiller, & Gazdar, 2007).

Interestingly, while the rates of tobacco smoking are decreasing in much of the Western world, there is no corresponding decline in the number of people

diagnosed with NSCLC (Jemal, et al., 2010). There is also an increasing epidemic in Japan, where lung cancer has been the leading cause of death since 1998 (Yano, et al., 2011). Importantly, the clinical manifestation of NSCLC in non-smokers is present disproportionately more in women than men suggesting that there are several strong predisposing characteristics, including female sex and being of Asian descent that contribute to an individual's chance of developing the disease (Provencio, Garcia-Campelo, Isla, & de Castro, 2009).

1.3.1 Molecular biology of Non-small-cell Lung Cancer

The development and progression of NSCLC is a multi-step process, characterised by the accumulation of multiple genetic and epigenetic lesions, which result in aberrant cell regulatory systems and growth-control pathways (Choi et al., 2007; Kolch & Pitt, 2010). Specifically, it has been shown that NSCLC tumour cells acquire common phenotypes such as the self-sufficiency of growth signals and resistance to anti-proliferative and apoptotic signals due to these molecular lesions (Mitsudomi, 2010).

Indeed, over the past couple of decades, the advent of next-generation molecular techniques and our ability to sequence entire genomes has uncovered the specific genetic alterations that confer uncontrolled proliferation and apoptotic resistance in NSCLC tumour cells (Bunn, Soriano, Johnson, & Heasley, 2000). This has not only significantly contributed to our understanding of the biological mechanisms that underlie tumourigenesis in NSCLC, but has also elucidated several key differences in this process in smokers and non-smokers (Ramalingam, Owonikoko, & Khuri, 2011).

Consequently, the evaluation of the pathogenesis and biological differences between NSCLC in smokers and non-smokers has contributed to a greater understanding of molecular differences which contribute to the development of the disease in these two groups. Accordingly, it is now hypothesised that NSCLC in smokers and non-smokers should be considered two different diseases based on the differences in the molecular lesions that contribute to their development (Sun, et al., 2007; Yano, et al., 2011).

1.3.2 Somatic Mutations in Non-small-cell Lung Cancer

Somatic mutations in cellular DNA underlie almost all human cancers (Weir et al., 2007). NSCLC is no different, as shown by Lee et al (2010) who sequenced the genome of an adenocarcinoma of a former smoker, and identified more than 50,000 single nucleotide variants and 530 somatic mutations, with an estimated genome-wide somatic mutation rate of 17.1 per megabase.

While the significance of the vast majority of mutations acquired by NSCLC tumours currently remains unknown, research has revealed that a small group of genetic lesions are not only necessary for the initial development or progression of the disease but also are required for the maintenance of that tumour's survival (Pao, Iafrate, & Su, 2011; Ramalingam, et al., 2011). Not surprisingly, these 'driving' oncogenic mutations are present in genes that encode proteins involved in key cellular pathways, and oncogenic activation of these proteins results in aberrant activity, thus contributing to the development of the hallmarks of cancer (Hanahan & Weinberg, 2000).

Specifically, two key proto-oncogenes have been identified in NSCLC, whereby the accumulation of a mutation within either of these genes is considered an important early step in driving pathogenesis and development of tumours (Pao, et al., 2011). These two genes are epidermal growth factor receptor (*EGFR*) and v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*). Both these genes encode proteins involved in cellular signalling pathways that control key characteristics of cellular growth and survival (Herbst, et al., 2008; Schubbert, Shannon, & Bollag, 2007).

1.4 Epidermal growth factor receptor (EGFR)

The epidermal growth factor receptor (EGFR) is one of four receptor tyrosine kinases of the ERBB or HER family of receptor kinases (Pao, et al., 2011; Sharma, et al., 2007). The *EGFR* gene is 188,306 base pairs (bp) and encodes a transmembrane glycoprotein which acts as a cell surface receptor that is routinely expressed in various tissues of epithelial, mesenchymal and neural origin (Zhang, Stiegler, Boggon, Kobayashi, & Halmos, 2010).

The protein encoded by the *EGFR* gene is a 170 kilo Daltons (kDa) cell membrane receptor, composed of an extracellular cysteine-rich ligand binding domain, a single alpha-helix transmembrane domain, and an intracellular domain with tyrosine kinase activity in the carboxyl-terminal tail (Figure 1) (Jimeno & Hidalgo, 2006; Kosaka et al., 2004; Takeuchi & Ito, 2010).

The crucial role of EGFR in normal development has been illustrated by a series of knockout mice models. Indeed, depending on their genetic background, mice without *EGFR* die between day 11.5 of gestation and day 20 after birth. Further, knockout mice show placental defects, lung immaturity, and various abnormalities in the bone, brain, heart, and various organs such as gastrointestinal tract, skin, hair follicles and eyes (Zhang, et al., 2010).

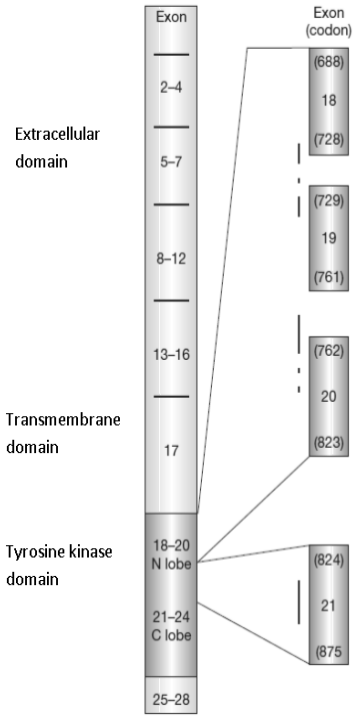

Figure 1 Schematic diagram showing the structure of the *EGFR* gene and the protein domains it encodes. There are three main domains: extracellular cysteine-rich domain (exons 2-16), transmembrane region (exon 17) and the tyrosine kinase domain. The tyrosine kinase domain encoded by exons 18 – 24 is magnified out to the right of this picture. The activating mutations are present in exon 19 and 21. Adapted from Abidoye et al. (2007)

1.4.1 EGFR cellular pathway

The EGFR cellular pathway was discovered by Stanley Cohen in the 1960s and since has become one of the highly characterised cellular pathways (Hanahan & Weinberg, 2000).The binding of epidermal growth factor, or another ligand, to the extracellular domain of EGFR induces homo-dimerisation or hetero-dimerisation with other members of ERBB family (Ansari, Palmer, Rea, & Hussain, 2009). The formation of these dimers results in conformational changes which activate the receptor and induce the tyrosine kinase catalytic activity the intracellular tyrosine kinase domain (Pao, et al., 2011).

Subsequently, this activation leads to auto-phosphorylation of one or more of the five tyrosine residues in the carboxyl-terminal tail, producing phosphotyrosine sites where adaptor and docking molecules will ultimately bind (Jimeno & Hidalgo, 2006). As such, the activation of EGFR ultimately leads to the phosphorylation of specific tyrosine residues in key members of several intracellular signalling pathways which promote cell proliferation and/or survival.

As shown in Figure 2, important intracellular pathways activated by EGFR include the PI3K-AKT-mTOR pathway involved in cell survival processes, and the RAS-RAF-MEK-ERK pathway involved in the transcription of molecules involved in cell proliferation, transformation and metastasis development (Franklin et al., 2010; Herbst, et al., 2008)

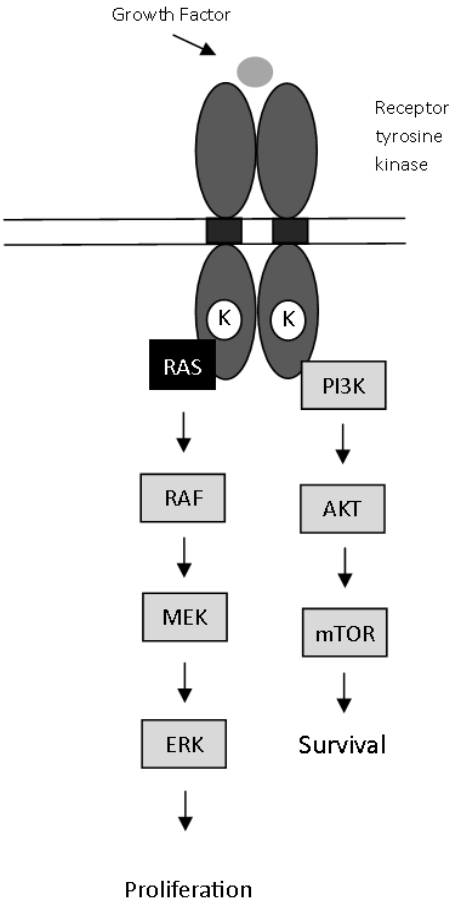

Figure 2 Schematic diagram showing EGFR and the associated cellular pathways. The PI3K-AKT-mTOR pathway is involved in cell survival while the RAF-MEK-ERK pathway is involved in cellular proliferation.

1.4.2 EGFR dysregulation and carcinogenesis

Given its pivotal role in cellular growth and development, the dysregulation of the EGFR gene or protein was suggested to be involved in carcinogenesis. Indeed, this was confirmed in the 1980s, when EGFR was demonstrated to be involved in tumourigenesis, and the overexpression of *EGFR* was shown to be associated with advanced stages of disease, resistance to conventional treatments and poor prognosis (Bearz, Berretta, Lleshi, & Tirelli, 2011). Since, *EGFR* has been found to be overexpressed in a variety of human malignancies including colon, breast, pancreas and NSCLC tumours (Kolch & Pitt, 2010). Interestingly, while more than 60% of NSCLC tumours show overexpression of *EGFR*, no overexpression is seen in SCLC (Zhang, et al., 2010).

Additionally, not only is the *EGFR* gene overexpressed in many cancers, but it has also been shown that the EGFR and its associated network is frequently altered in many cancers (Kolch & Pitt, 2010). This resultant aberrant activity of EGFR and the associated downstream signalling pathways have been shown to be involved in the development of several hallmarks of cancer including proliferation, survival, invasiveness, metastatic spread and tumour angiogenesis (Hanahan & Weinberg, 2000; Herbst, et al., 2008).

1.4.3 EGFR activating mutations

In 2004, two separate groups of investigators independently identified somatic mutations in the *EGFR* gene in NSCLC adenocarcinomas (Ladanyi & Pao, 2008). These mutations were found in approximately 10% of tumours from patients in the United States and in 30-50% of tumours from patients in Asia (Herbst, et al., 2008; Pao, et al., 2011).

Approximately 90% of lung-cancer-specific *EGFR* mutations were shown to affect a few specific amino acid residues. These gain of function mutations comprise either a leucine-to-arginine substitution at codon position 858 (L858R) in exon 21 or a deletion in exon 19 that affects the conserved sequence, leucine-arginine-glutamic acid-alanine motif (LREA), centred around codons 746-750 (delE746-A750) (Dacic, Shuai, Yousem, Ohori, & Nikiforova, 2010; Pao, et al., 2011; R. Rosell et al., 2009). The in-frame deletions in exon 19 make up 45% to 50% of mutations, while the missense mutation in exon 21 makes up 35% to 45% (Sequist, Bell, Lynch, & Haber, 2007). Both of these mutations are present in the part of the gene that encodes the tyrosine kinase (TK) domain of EGFR (Figure 3).

Significantly, investigators found that regardless of ethnicity these mutations are present in NSCLC tumours of never smokers (defined as smoked less than 100 cigarettes in a patient's lifetime), female sex, and with adenocarcinoma histology (Pao, et al., 2011). Furthermore, the frequency of *EGFR* mutations is inversely associated with cumulative smoking exposure suggesting a strong correlation *EGFR* mutations and the development of lung cancer in non-smokers (Yano, et al., 2011). Several other mutations in the TK domain of

EGFR have been reported (Appendix I), yet the clinical significance of these rare mutations has yet to be determined (Pao & Girard, 2010).

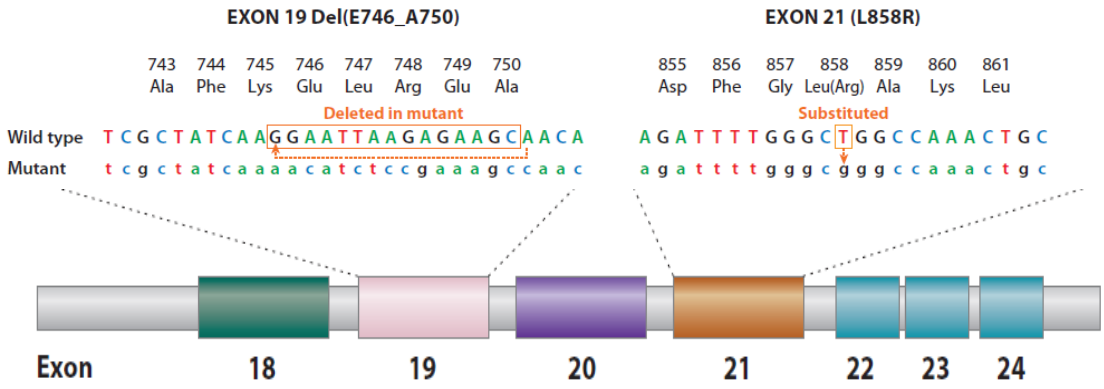

Figure 3 EGFR tyrosine kinase domain (Exons 18 – Exon 24) showing LREA deletion in Exon 19 and missense mutation L858R in Exon 21. Adapted from Yano et al. (2011)

1.4.4 Role of EGFR activating mutations in NSCLC

EGFR mutations in the tyrosine kinase (TK) domain of the receptor are considered gain of function, or activating mutations, as the protein produced from a mutated *EGFR* gene has a constituently active TK activity (Kosaka, et al., 2004). It has been shown that these mutations disrupt the auto-inhibitory interactions within the TK domain and consequently cause the induction of higher phosphorylation of EGFR compared with the wild-type protein (Yatabe & Mitsudomi, 2007).

Whilst under physiological conditions ligand binding is required for the activation of EGFR, these mutations enable independent receptor activation (Sharma, et al., 2007). This ligand-independent activation EGFR results in constitutive signal transduction and an increase in downstream signalling. Indeed, crystal structure analyses of wild-type and L858R EGFR have shown that the substitution of lysine with arginine results in a much larger side chain, which locks the protein in a constitutively active state (Pao, et al., 2011). Whereas the LREA deletion results in an altered amino-acid sequence adjacent to the ATP binding cleft of EGFR mutant, conferring not only ligand-dependant activation, but also increasing the duration and activation of receptor

signalling after ligand binding compared with wild-type receptor (Rafael Rosell et al., 2010).

As such, through the activation of increased EGFR signalling, these mutations confer a selective advantage for these cells, to which the malignant cells become dependent. Indeed, the tumour cells require the increased EGFR-mediated anti-apoptotic signalling merely to sustain survival and are highly sensitive to any interruption of these pathways. As a consequence, in developing their malignancy, these cells become more dependent on the oncogenic activation of EGFR and thus provide an ideal target for anti-cancer therapy (discussed further in sections 1.7.1 and 1.8.2) (Yoshida, Zhang, & Haura, 2010).

1.5 v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS)

The *KRAS* gene encodes a small quinine nucleotide transferase molecule that links cell surface receptors, including EGFR, to their associated intracellular pathways (Franklin, et al., 2010). KRAS is a member of the Ras superfamily which are characterised by the presence of a catalytic G domain and play a crucial role in the transduction extra- and intra-cellular signalling (Jančík, Drábek, Radzioch, & Hajdúch, 2010).

The product of the *KRAS* gene is a 21.6kDa protein composed of 188 amino acid residues. The KRAS protein plays a key role in the kinase cascades such as RAS/MAPK/ERK, PI3K/AKT/STAT3 and STAT5 signalling pathways involved in a number of cellular processes including proliferation, differentiation and apoptosis (O'Hagan & Heyer, 2011).

1.5.1 KRAS role in EGFR cellular pathway

KRAS is a key member of the EGFR pathway, and can be thought of as a signal switch molecule that couples the EGFR receptor activation to downstream effector pathways (Schubbert, et al., 2007). Indeed, KRAS cycles between active guanosine triphosphate (GTP)-bound and inactive guanosine diphosphate (GDP)-bound forms whereby the active form of the protein binds to more than 20 effector proteins and stimulates downstream signalling cascades (Suda, Tomizawa, & Mitsudomi, 2010).

Additionally, the KRAS protein has intrinsic GTPase activity, which plays an important role in signal transduction through multiple downstream transducers, or GTPase-activating proteins (GAPs), which amplify the GTPase activity of the KRAS protein 100,000-fold (Jančík, et al., 2010). Accordingly, as seen in Figure 2, KRAS is an important early downstream mediator of EGFR-induced cell signalling (Rafael Rosell, et al., 2010).

1.5.2 *KRAS* activating mutations

KRAS was first identified as an oncogene in 1982 and is now recognised as one of the most frequently mutated oncogenes in human cancer (Pritchard, Akagi,

Reddy, Joseph, & Tait, 2010). Accordingly, activating *KRAS* mutations are commonly present in epithelial malignancies, accounting for about 30% of all human cancers (O'Hagan & Heyer, 2011). Significantly, *KRAS* mutations are prevalent in pancreatic, colorectal, endometrial, biliary tact, cervical and lung cancers (Schubbert, et al., 2007).

Indeed, tumours harbouring *KRAS* mutations are often characterised as highly aggressive cancers, and consequently the presence of these mutations correlated with a poor prognosis in NSCLC (Bunn, et al., 2000; Pritchard, et al., 2010; Schubbert, et al., 2007). More than 95% of lung-cancer-specific *KRAS* mutations affect two specific amino acids in codon 12 or 13 in exon 2 of the gene (Table 1). These gain of function mutations comprise single-point substitutions that change the glycine encoded at both these codons (Pritchard, et al., 2010).

These oncogenic point mutations impair the intrinsic GTPase activity of the KRAS protein and confer resistance to GAPs thereby causing the protein to accumulate in its active GTP-bound state (Franklin, et al., 2010). As such, KRAS proteins encoded by a mutated *KRAS* gene have lost the intrinsic negative-feedback control of their activity, leaving the mutated protein constitutively activated, resulting in the dysregulation of the associated key cellular pathways (Jančík, et al., 2010).

Table 1 Activating missense *KRAS* mutations in NSCLC. Adapted from Kristensen et al. (2010)

Exon	Codon	Nucleotide change	Amino acid change
2	12 (GGT)	c.35G>C	p.Gly12Ala
		c.35G>A	p.Gly12Asp
		c.34G>C	p.Gly12Arg
		c.34G>T	p.Gly12Cys
		c.34G>A	p.Gly12Ser
		c.35G>T	p.Gly12Val
	13 (GGC)	c.38G>A	p.Gly13Asp

1.5.3 Role of *KRAS* activating mutations in NSCLC

KRAS oncogenic mutations activate the KRAS protein independently of ligand binding to EGFR, and subsequently enable the transmission of the associated signalling pathways independently of EGFR (Kristensen et al., 2010). Specifically, activating *KRAS* mutations led to activation of the RAS/MAPK pathway, inducing multiple cellular processes including proliferation and differentiation leading to tumour growth (Zuo et al., 2009).

While the reported rate of *KRAS* mutations in NSCLC patients varies significantly with ethnicity, their presence is routinely concomitant with a history of tobacco smoking and poor prognosis (Bearz, Berretta, Lleshi, & Tirelli, 2011; Provencio, et al., 2009). As such they occur predominately in Caucasian patients with an incidence rate of about 33% whereas they are only present in 10% of East Asian patients (H. Do, Krypuy, Mitchell, Fox, & Dobrovic, 2008; Kristensen, et al., 2010). Further, *KRAS* mutations are found predominately in the adenocarcinoma histological subtype of NSCLC, rarely in squamous cell carcinoma and never in SCLC (Borras et al., 2011; Suda, et al., 2010).

Genetically engineered animal models have provided invaluable information about the role of activating mutations in NSCLC (O'Hagan & Heyer, 2011). Indeed, nicotine-derived nitroaminoketone (NNK), a potent tobacco-specific carcinogen, has been shown to readily induce *KRAS* missense mutations in rodents especially at codon 12, further highlighting the significant association between smoking and KRAS mutations (Borras, et al., 2011; Sun, et al., 2007).

1.6 Clinical Significance of *EGFR* and *KRAS* mutations in NSCLC

While NSCLC tumours harbouring *EGFR* or *KRAS* mutations look the same under a microscope, their underlying genetic and molecular makeup is significantly different. Given that both *EGFR* and *KRAS* mutations have been shown to play an important role in the pathogenesis of NSCLC adenocarcinomas and have a strictly mutually exclusive relationship, it has been proposed that they have functionally equivalent roles in tumourigenesis (Kosaka, et al., 2004; Pao et al., 2005).

This has led to the hypothesis that the mutation status of these two oncogenes defines the two clinically relevant molecular subsets of NSCLC, in that *EGFR* mutations are vital to the pathogenesis of non-smoking NSCLC while *KRAS* mutations are limited to driving the smoking-associated NSCLC. The emergence of these two genotypically different tumours has contributed to a paradigm shift, in that the molecular structure and function of NSCLC tumours is considered to be more useful than histological type in predicting the aggressiveness, sensitivity to therapy and prognosis (as presented in Table 2) (Works & Gallucci, 1996).

The emergence of clinical and epidemiological studies focused on the biological and genetic difference between smoking- and non-smoking associated NSCLC and their associated risk factors, has provided further evidence for the need to classify NSCLC tumours on the basis of the mutational status of these two oncogenes. Indeed, elucidating these differences has provided researchers with promising new anti-cancer therapies which have been developed with the goal of longer survival or an amelioration of symptoms (Herbst, et al., 2008; Nose, Uramoto, Iwata, Hanagiri, & Yasumoto, 2011).

1.6.1 EGFR and KRAS as oncogenic ‘driver’ mutations

EGFR and *KRAS* mutations are considered ‘driver’ mutations in NSCLC as their presence is thought to be crucial not only for the establishment, but also the viability of the tumour (Pao, et al., 2011). In an experiment which compared tumour tissue with matched normal tissue from NSCLC patients, *KRAS* and *EGFR* where shown to arise somatically during tumour formation, providing further evidence for the importance of these mutations as an early

oncogenic event in the development of NSCLC (Jimeno & Hidalgo, 2006). Indeed, lung cancer can be induced on mice through the generation of *KRAS* exon 12 mutations and the subsequent suppression of this mutated *KRAS* gene in these mice resulted in the reduction and elimination of these tumours (Borras, et al., 2011).

As such, accumulating evidence suggests that these two somatic mutations are important oncogenic events in the development and maintenance of the malignant phenotype in NSCLC (Yatabe & Mitsudomi, 2007). Consequently, *EGFR* and *KRAS* mutations are considered to drive carcinogenesis in the two disease subtypes, either through the non-smoking EGFR mediated pathway or the smoking-associated activated KRAS pathway, respectively.

Table 2 Comparsion of *KRAS* and E*GFR* mutations in NSCLC highlighting the key contrasting features. Adapted from Kristensen et al. (2010)

	KRAS	**EGFR**
Discovery of mutation	1982	2004
Biochemical function	Small GTP-binding protein	Receptor tyrosine kinase
Common mutations	Missense mutation at codons 12 or 13	Exon 19 LREA deletion, missense mutation at codon 858 (L858R) in Exon 21
Mutation effect	Impaired GTPase-activity resulting in constitutive activation	Loss of auto-inhibitory feedback in tyrosine kinase domain
Mutation in tumours other than NSCLC	Common – pancreas, colon, bile duct etc.	Rare
Smoking status	Smokers	Non-smokers
Ethnicity	Caucasians > East Asians	East Asians > Caucasians
Sex	Male > Female	Female > Male
Histology	Adenocarcinoma	Adenocarcinoma
Prognostic impact	Poorer	Better
Response Rate for EGFR-TKI therapy	0%	70-90%

1.7 Targeted therapies in NSCLC

Targeted therapies are drugs or other substances that block the growth and spread of cancer by interfering with specific molecules involved in tumourigenesis and progression (Squassina et al., 2010). Given the high mortality rate associated with NSCLC, current treatment paradigms are shifting from cytotoxic chemotherapy to molecular targeted therapies in the aim of reducing side-effects and circumventing the therapeutic plateau reached by traditional chemotherapy (Pao, et al., 2011).

Accordingly, by selectively targeting cancer-specific characteristics, targeted therapies may be more effective than chemotherapy and less harmful to normal cells (Gazdar, 2010). Indeed, targeted therapies have provided some improvement in clinical outcomes and as a consequence there has been an explosion in number of targeted therapies employed for the therapeutic treatment of many cancers, including NSCLC (Hirsch, 2009).

1.7.1 EGFR Tyrosine Kinase Inhibitors

Given the biological importance of EGFR and its associated molecular pathways, several molecules have been developed to inhibit EGFR in attempt to block EGFR signalling. As such, EGFR antibodies have been successfully introduced for the treatment of colon cancer, while EGFR specific inhibitors have been introduced for the treatment of lung cancer (Ansari, et al., 2009; Mancl, Kolesar, & Vermeulen, 2009).

EGFR tyrosine kinase inhibitors (TKIs) specifically inhibit the tyrosine kinase (TK) domain of EGFR and have been investigated as first-line or subsequent therapeutic options for patients with advanced NSCLC (Gazdar, 2010). These EGFR-TKIs are orally active, small molecular weight quinazolinamine derivatives (Figure 3) (Jimeno & Hidalgo, 2006; Pao, et al., 2005). They selectively inhibit EGFR-mediated cellular signalling by binding to the adenosine triphosphate (ATP) binding pocket of the TK domain of EGFR, and consequently block downstream signalling (Nomoto et al., 2006).

As such, EGFR-TKIs cause apoptosis in cancer cells due to the shutdown of PI3K-Akt and ERK1/2 signalling pathways following the inhibition of EGFR activation (Takeuchi & Ito, 2010). The rationale behind TKIs is that by

targeting EGFR, they can selectively target cancer cells with an oncogenically transformed EGFR, given that these cells are more dependent on EGFR-mediated signalling for survival and homeostasis. Indeed, there are only minor common side effects of EGFR-TKIs, including the development of a rash and acne or diarrhoea, and unlike traditional cytotoxic agents they do not cause myelosuppression, neuropathy, alopecia and severe nausea (Mok, 2009).

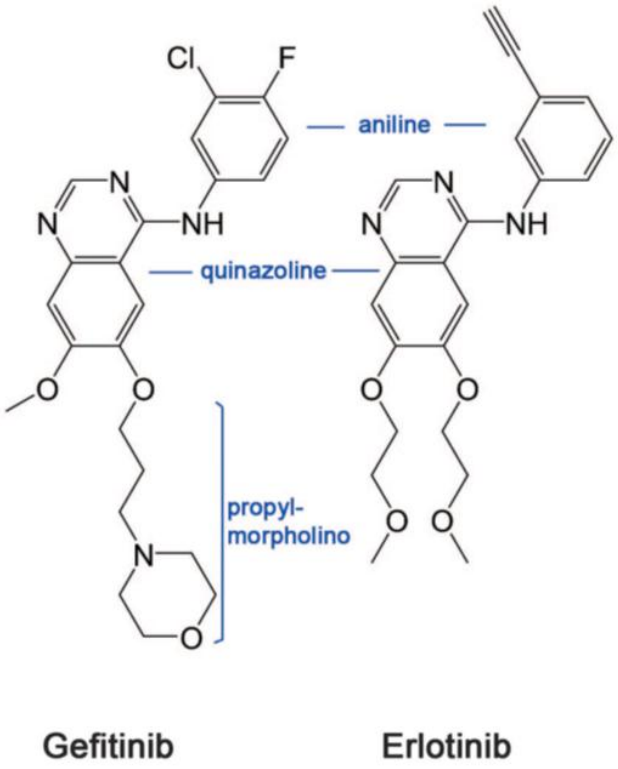

Figure 4 Chemical structures of EGFR-TKIs gefitinib and erlotinib. These EGFR-TKIs act as competitive inhibitors by competing with ATP for the ATP-binding site of the TK domain of the EGFR receptor. Adapted from Mok (2009)

1.7.2 Gefitinib and Erlotinib

There are currently two EGFR specific TKIs, gefitinib and erlotinib, employed in the treatment of advanced NSCLC (Santarpia et al., 2011). Gefitinib (Iressa) was the first to be approved by the US Food and Drug Administration (FDA) in 2003 (Allison, 2010). Erlotinib (Tarceva) was approved in 2004 by the FDA and the European Medicines Agency in 2005 as a second-line therapy for those NSCLC patients with advanced or metastatic tumours (Penzel et al., 2011).

Both these drugs have been shown to reversibly and selectively target EGFR and block signal transduction with minimal activity against other tyrosine and serine/threonine kinases (Masago, et al., 2008). Current clinical practice in most of the world, including NZ, is to use EGFR-TKIs as second- or third-line therapy in advanced NSCLC (Sequist & Lynch, 2008). As of early 2011, gefitinib is licensed in 66 countries for the treatment of NSCLC (Cataldo, et al., 2011).

1.8 Patient Response to EGFR-TKIs

EGFR-TKIs have demonstrated activity in advanced NSCLC in the second- and third-line setting in several Phase II and III clinical trials (as summarised in Appendix II) (Hann & Brahmer, 2007). However, in contrast to conventional chemotherapy, EGFR-TKIs appear to be more effective in specific NSCLC populations. Indeed, subset analysis of these several large scale clinical trials has shown that never smokers, people of Asian descent, patients with adenocarcinoma histology and females are more likely to respond to EGFR-TKIs (Fukui et al., 2008; Wu et al., 2010; Yatabe & Mitsudomi, 2007).

Specifically, early Phase II trials demonstrated that EGFR TKI monotherapy has modest activity in the general NSCLC population, with response rates of about 12-26% (Cappuzzo et al., 2005; Hann & Brahmer, 2007). However, subsequent Phase III trials reported contrasting results, in that tumours in the vast majority of patients failed to respond, while a minority showed dramatic tumour shrinkage and symptomatic improvement (T.-Y. Kim, Han, & Bang, 2007).

1.8.1 Retrospective analysis of Clinical Trials for EGFR-TKIs in NSCLC

Given the variable response to TKIs in the NSCLC population, retrospective analysis of Phase II and III clinical trials were undertaken to explain the differential therapeutic response shown by certain subsets of NSCLC patients (Yatabe & Mitsudomi, 2007). Consequently, the differential response shown by NSCLC patients was attributed to the genetic makeup of their respective tumours (Jimeno & Hidalgo, 2006).

Indeed, while it had already been established that never smokers, females, people of Asian descent, and patients with adenocarcinoma histology showed therapeutic benefit to TKIs, this correlation was shown to be the result of these patients harbouring EGFR-mutant tumours (Pao, et al., 2011; Sequist & Lynch, 2008).

1.8.2 EGFR activating mutations and hypersensitivity to EGFR-TKIs

Since their identification in 2004, EGFR activating mutations have shown a strong association with positive clinical outcomes in parameters such as response rate, time to progression, and overall survival in those patients treated with TKIs (Nomoto, et al., 2006). Given that EGFR-TKIs interact with this binding cleft in the EGFR protein, it is thought that these oncogenic mutations confer greater sensitivity to TKIs as these drugs to bind more tightly to the mutated protein than to the wild-type (Sequist & Lynch, 2008).

As such, not only do cells with activated EGFR have a better affinity for the binding of these competitive inhibitors, but have oncogenic addiction, in that they are dependent on EGFR signalling for cellular growth and proliferation (C. H. Lin, et al., 2010). As such, it is thought that the mutated EGFR is the Achilles heel of these tumours, as the tumour cannot maintain a malignant phenotype when EGFR-mediated signalling is interrupted by EGFR-TKIs.

1.8.3 EGFR secondary mutations and acquired resistance to TKIs

Unfortunately, virtually all patients who initially respond to EGFR-TKIs eventually develop acquired resistance throughout the course of their treatment (Ercan et al., 2010). Studies have shown in approximately 50% of these cases that this resistance is due to a secondary mutation, T790M in exon 20 of the *EGFR* gene (Borras, et al., 2011; Provencio, et al., 2009; R. Rosell et al., 2005).

The EGFR T790M mutation is oncogenic by itself, but when present in conjunction with another *EGFR* activating mutation, the double mutant leads to a substantial increase in EGFR signalling and oncogenic transformation both *in vivo* and *in vitro* (Ercan, et al., 2010). As such, the T790M mutation is associated with a short median progression free survival (PFS) of only 7.7 months in patients carrying the mutation compared with 16.5 months in those without the mutation (Rafael Rosell, et al., 2010).

1.8.3.1 Mechanism of T790M acquired resistance to EGFR-TKIs

Biochemical studies combined with the molecular analysis of tumour material from patients whom have developed resistance have shown that the T790M mutation is a common cause of acquired resistance in NSCLC patients. The threonine at position 790 is the 'gatekeeper' residue as its location at the entrance of the hydrophobic pocket of the ATP binding cleft means it has important role in determining inhibitor specificity in the receptor (Yun et al., 2008).

As such, Yun et al (2008) carried out enzyme kinetic assays using an ATP/NADH coupled assay system whereby they could detect the rate of ATP hydrolysis in both the wild-type and T790M mutant. The kinetic characterisation of the WT and the mutant EGFR kinases revealed a marked decrease in the Michealis-Menton constant (K_m) for ATP in the drug resistant T790M mutant compared with the drug sensitive L858R mutant. Consequently, it is now thought that the T790M restores the receptors affinity for ATP, enabling ATP to outcompete TKIs for the ATP binding cleft of the tyrosine kinase domain, re-establishing signalling from EGFR.

1.9 *EGFR* and *KRAS* mutations as biomarkers for response to TKIs

Biomarkers can be used to describe any useful characteristic that can be measured and used as an indicator of a normal biological process, a pathological process, a pathogenic process or in the case of *EGFR* and *KRAS* activating mutations, a pharmacologic response to a therapeutic agent, such as EGFR-TKIs (Squassina, et al., 2010). The discovery of the association of *EGFR* and *KRAS* mutations with sensitivity and resistance, respectively, to EGFR-TKIs represents a major breakthrough in the application of targeted therapies in NSCLC (van Zandwijk et al., 2007).

Specifically, therapeutic benefit has been seen in approximately 75% of patients whose tumours harbour drug-sensitising *EGFR* mutations, compared with <5% of NSCLC patients with wild-type receptor and <1% of patients with *KRAS* mutant tumours (Pao, et al., 2009). Given this significant association, the detection of these predictive biomarkers provides physicians with the opportunity to tailor personalised treatment for NSCLC patients.

1.10 Personalised Medicine in NSCLC

Personalised medicine focuses on the individualised drug treatment of diseases according to each patient's underlying molecular and genetic makeup (Xie & Frueh, 2005). Specifically, it refers to the delivery of the right drug to the right patient at the right dose. As such, the application of genetic information in order to develop targeted strategies is known as pharmacogenomics, whereby those individuals likely to respond are identified on the basis of a wide range of data, the most important of which are environmental and genetic factors (Squassina, et al., 2010).

Given the significant variability in patient response to EGFR-TKIs, the development of treatment decisions based on the expression of genetic biomarkers facilitates the translation from research to clinical usefulness, as well as the establishment of more effective and cost-saving pathways for patient treatment. As such, the goal of personalised medicine is to maximise the likelihood of therapeutic benefit and minimise adverse reactions, in attempt to optimise patient treatment. This is of particular importance in NSCLC given the modest

improvements in quality of life and overall survival with traditional chemotherapy, the high mortality and prevalence of the disease, and the burden that patient management places on the public health system (Triano, et al., 2010).

1.11 Importance of developing clinically relevant assay to detect *EGFR* and *KRAS* mutations

The advent of personalised medicine represents a paradigm shift in the treatment of NSCLC, whereby the genetic and molecular characterisation of a patients tumour is more important than the histological typing to guide treatment regimens and predict patient response (Cho, 2007). As such, gaining knowledge of the genetic and molecular characteristics of an individual patient's tumour has become an essential step for treatment decision-making in NSCLC (Luthra & Zuo, 2009).

Indeed, given the growing evidence that the mutation profile of individual tumours is highly predictive of patient response to EGFR-TKIs, they must be analysed in specific populations (Pierre-Jean & William, 2012). As such, the establishment of a clinically relevant assay to detect *EGFR* and *KRAS* mutations in NSCLC patients is crucial for the translation of genetic information in a clinical setting to aid decision making and personalised treatment. Given that clinical trials have shown response rates are as low as 12% in unselected patients receiving EGFR-TKIs, and one round of treatment with gefitinib costs over US$10,000, there is significant economic value in pre-screening patients for these predictive biomarkers in NSCLC patients (Cappuzzo, et al., 2005; Dacic, et al., 2010).

Due to the highly predictive nature of *EGFR* and *KRAS* mutations for patient response, or lack thereof, to EGFR-TKIs, the establishment of an accurate, rapid and inexpensive assay to stratify patients on the basis of the mutation status of their tumours would enable the selection of optimal therapy and ultimately improve patient outcomes.

1.11.1 Obstacles in development of clinically relevant assay to detect *EGFR* and *KRAS* mutations

The routine molecular classification of NSCLC tumours in clinical practice requires a diagnostic assay which is accurate, rapid, cost effective, and requires little effort to optimise, perform and analyse (Milbury, Li, & Makrigiorgos, 2009). As such, there is a significant need to establish of a rapid and robust assay for the detection of *EGFR* and *KRAS* mutations in NSCLC patient tumours (Krypuy, Newnham, Thomas, Conron, & Dobrovic, 2006).

However, there are many obstacles that limit the sensitivity and selectivity of current diagnostic techniques, which has contributed to the lack of guidelines for mutation testing and other DNA analysis in clinical practice (Squassina, et al., 2010). Indeed, there are several constraints placed on the establishment of a clinically relevant assay to detect *EGFR* and *KRAS* mutations in NSCLC patients, the most important of which will be discussed in the following sections.

1.11.1.1 Biopsy Specimens

As surgical resection is not a viable for most NSCLC patients, biopsy specimens of patient tumours are taken via fine-needle aspiration. Fine needle aspiration biopsy, while reducing the discomfort for the patient, creates potential problems for genetic testing of specimens. Indeed, while these fine-needle aspirates provide sufficient tumour tissue for histological classification by a pathologist, the limited tissue can result in a number of false negatives in genetic testing, whereby the presence of a mutation can go undetected (Sequist, et al., 2007).

As such, these tumour specimens tend to vary in size, not only in the overall specimen size, but also in the ratio of tumour to normal tissue, which further limits the sensitivity of screening procedures due to the lack of mutated alleles and/or the background of wild-type alleles (da Cunha Santos, Saieg, Geddie, & Leighl, 2011). Consequently, some researchers have suggested that the small number of EGFR mutation positive results seen in NSCLC patients has limited the statistical significance of the association of these mutations with drug

response, hindering the application of genetic information in order to stratify patients and develop targeted therapies (Squassina, et al., 2010).

1.11.1.2 FFPE Specimens

Biological specimens, including those taken from NSCLC patients, are fixed using formalin and embedded in paraffin wax for the use in clinical pathology and long-term storage. Indeed, formalin-fixed, paraffin embedded (FFPE) specimens are an invaluable source of genomic DNA and have been extensively utilised in genetic studies and retrospective analysis of diseases (Furuta et al., 2006). However, the chemical processes involved in the fixation of biological specimens hinder the diagnostic techniques involved in the development of our understanding at the molecular level.

Specifically, chemical reactions occur between formaldehyde, the active component of formalin, and the genomic DNA within the specimen. Such reactions include methylol formation, methylene bridge formation, apurinic and apyrmidinic site formation and hydrolysis of phosphodiester bonds (Hongdo Do & Dobrovic, 2009). These reactions cause the DNA and other nucleic acids to become degraded and fragmented which is problematic during PCR thermocycling as it creates "blocks" for the DNA polymerase (Farrugia, Keyser, & Ludes, 2010).

Indeed, as PCR is a universal first step utilised by most molecular diagnostic techniques, this limits the sensitivity of almost all established procedures. Accordingly, the failure of PCR using FFPE specimens due to the generation of DNA-protein cross-linkages and nucleic acid fragmentation is a considerable obstacle it the establishment of a clinically relevant assay. Steps have been made to combat this, which include the design of primer sets that give PCR products less than 300base pairs (bp), the use of high-fidelity DNA *Taq* polymerase with proofreading and exonuclease activity and the extensive optimisation PCR conditions for amplification.

1.11.1.3 DNA extraction from FFPE specimens

The DNA-protein cross-linkages and nucleic acid fragmentation as are result of FFPE also complicates the isolation of quality genomic DNA from these specimens (Farrugia, et al., 2010). In attempt to combat this, a number of different protocols have been developed for the extraction of DNA from FFPE specimens, which vary in terms of their complication, turnaround time and reported degrees of success (Siwoski et al., 2002). Several studies have been conducted to compare the efficacy and suitability of these for clinical assays, yet given the complicated and expensive reagents used in the majority of these methods, none have been established for the routine use on a large scale clinical setting.

Such methods include the phenol-chloroform method, xylene/ethanol method, ammonium acetate precipitation, mineral oil, antigen retrieval principle (heating under influence of pH), and commercially available kits based on silica binding principle and other methods are becoming readily available (Farrugia et al 2010).

1.11.1.4 Direct Sequencing

Direct sequencing is considered the 'gold standard' for mutations screening and is commonly employed for genotyping biological specimens. However, direct sequencing has a sensitivity limit of 20%, whereby it is unable to detect mutations present in a specimen below this frequency (Morlan, Baker, & Sinicropi, 2009; Uhara et al., 2009). As such, recent studies have shown that the use of direct sequencing for *EGFR* and *KRAS* mutation detection in FFPE NSCLC specimens results in a number of false-negatives, whereby the presence of these oncogenic mutations will go undetected.

Indeed, due to the highly heterogeneous nature of NSCLC tumour specimens, and biopsy specimens being contaminated with normal cells, these clinical relevant mutations are extensively diluted in wild-type alleles below the detection limit of conventional PCR-direct sequencing based assays, requiring the need for other DNA screening and diagnostic techniques (Kristensen, et al., 2010).

1.12 Purpose and Scope of Investigation

The mutation status of the two oncogenes, *EGFR* and *KRAS*, has shown noticeable variation in the NSCLC disease phenotypes present in either non-smokers or current and former smokers. Given the clinical significance of these mutations, the development and optimisation of PCR-based methods for their detection is crucial for the establishment of personalised patient management.

The primary focus of this investigation is the detection of *EGFR* and *KRAS* mutations in FFPE specimens. As such, this investigation will attempt to address the obstacles outlined above by investigating the use of a variety of genotyping and mutation detection protocols, in attempt to establish a clinically relevant *EGFR* and *KRAS* mutation detection assay for the use on FFPE NSCLC specimens.

As such, clinically relevant diagnostic assays must address the sensitivity limit of standard PCR and direct sequencing. In attempt to do so, several other methods have established for DNA diagnostics, with greater sensitivity and selectivity for mutation detection. However, these methods are limited, due to the requirement of increased post-PCR manipulation, cost of reagents and equipment and timely nature of these procedures, which is problematic for the clinical application of these assays.

1.13 Experimental methods and assay design

Given that PCR may fail to detect mutations when there is a high 'background' of wild-type DNA in a specimen, the optimisation of PCR conditions for the amplification of genomic DNA from biological specimens is crucial to increase the sensitivity of any downstream analysis. As such, several protocols will be employed to address the sensitivity limit of standard PCR and direct sequencing. These protocols will briefly be introduced in the subsequent paragraphs and summarised in Figure 5.

1.13.1 COLD-PCR

Co-amplification at Lower Denaturation temperature (COLD-PCR) is a modification of conventional PCR to selectively amplify mutant alleles in a wild type background (Zuo, et al., 2009). This method is based on the observation that a given DNA sequence has a critical denaturation temperature (T_c), which is lower than its melting temperature (T_m). This T_c is defined as the temperature below which PCR amplification efficiency for that sequence drops significantly (Li et al., 2008).

The principle behind COLD-PCR is that a single nucleotide mismatch anywhere along a double-stranded DNA sequence will generate a small but predictable change in the melting temperature (T_m) for that given sequence (Li & Makrigiorgos, 2009). Specifically, depending on the context and position of the mismatch, a mutation can change the T_m of a sequence by 0.2-1.5^oC in sequences up to 200bp long.

Indeed, PCR with the denaturation temperature, set to T_c during thermocycling contributes to the selective amplification of minority (mutated) alleles, as it results in mainly heteroduplexes (formed by the hybridisation of mutant and wild-type sequences) to be denatured and amplified, and leaves the wild-type homoduplexes double-stranded and not amplified efficiently (Milbury, et al., 2009; Zuo, et al., 2009). As such, COLD-PCR is of particular value, as changing a single parameter of thermocycling during PCR is a cost-effective way to increase the sensitivity of PCR-based methods for mutation detection.

1.13.2 Restriction Enzyme Digestion mediated mutant enrichment PCR

Restriction Enzyme digest-mediated mutant enrichment PCR is a two-step PCR method with intermittent restrictive digestion. The restrictive digestion selectively eliminates the wild-type genes, resulting in the enrichment of mutant alleles for the second round of amplification (He et al., 2009). As such, the mutant-enriched PCR is an assay with a high specificity and sensitivity that can detect one mutant gene among as many as 10^3 to 10^4 copies of the wild-type gene (Asano et al., 2006).

Simply, a first round of PCR is carried out on the genomic DNA isolated from the specimen and this product is then incubated with the restriction enzyme which will cut the DNA at its corresponding recognition site. An aliquot of this digestion product is then used as the template for a second round of PCR. During PCR of the digestion product, the mutant sequences will be selectively amplified, as the digested wild-type sequences will not be amplified with the same efficiency (Hlinkova, Babal, Berzinec, Majer, & Ilencikova, 2011a).

The amplification products from the second round of PCR are subjected to analysis via electrophoresis on a polyacrylamide gel. As such, mutation containing specimens will show a different banding pattern when compared to the control. Indeed, this method requires a fair amount of post-PCR processing which results in a time-consuming and labour intensive protocol. Further, genotyping via restriction digestion and PAGE gels can be confounded by a number of complicating factors including variation in DNA quantity and variability of silver staining.

1.13.3 Single-stranded Conformation Polymorphism

Single-stranded Conformation Polymorphism (SSCP) is a well-established method of detecting low level mutations in a background of wild-type alleles. Mutation detection by SSCP relies on the fact that when PCR product is denatured, the resultant single-stranded DNA, when cooled rapidly, adopts secondary conformations according to sequence (Jacobson, 1998). When a sample containing a mutated single stranded PCR product is electrophoresed on an acrylamide gel, the mutant DNA, because it has a different sequence, and hence conformation, compared to normal controls, will display altered migration.

As such, SSCP is of particular value because it can detect the presence of genetic aberrations which may not be detected by direct sequencing because of sequence masking by background normal sequences (Hu et al., 2007). Hence, SSCP is a simple, although only moderately sensitive, technique for screening PCR products for the presence of mutant sequences. However, due to the need for gel conformation SSCP requires a lot of post-PCR processing and handling,

this increases the turnaround time and the chance of contamination. Indeed, complications including inconsistent and time consuming silver staining of SSCP gels confound the interpretation of SSCP evidence.

1.13.4 DNA melt curves analysis

DNA melt curve analysis, in particular high resolution DNA melting (HRM) analysis, is a relatively new method for detecting mutations that are present in specimens with increased specificity and sensitivity compared with previously mentioned protocols. Indeed, it is reported that conventional PCR followed by HRM has a mutant detection limit of 2% (Fassina et al., 2009; Penzel, et al., 2011).

Mutation detection by DNA melt curve analysis relies on the fact that when a PCR product is heated in the presence of an intercalating dye, the melting profile of the sequence can be monitored. Indeed, the melting profile is dependent on the DNA sequence, and any mutations will result in the formation of heteroduplexes which have a different melt profile (Reed, Kent, & Wittwer, 2007).

As such, DNA melt curve analysis is simpler and more cost effective way to genotype specimens as it does not require post-PCR processing as analysis can be performed on the same instrument. This is highly crucial in clinical practice, not only is there reduced chance of contamination, but this will enable timely genetic information to aid decision making.

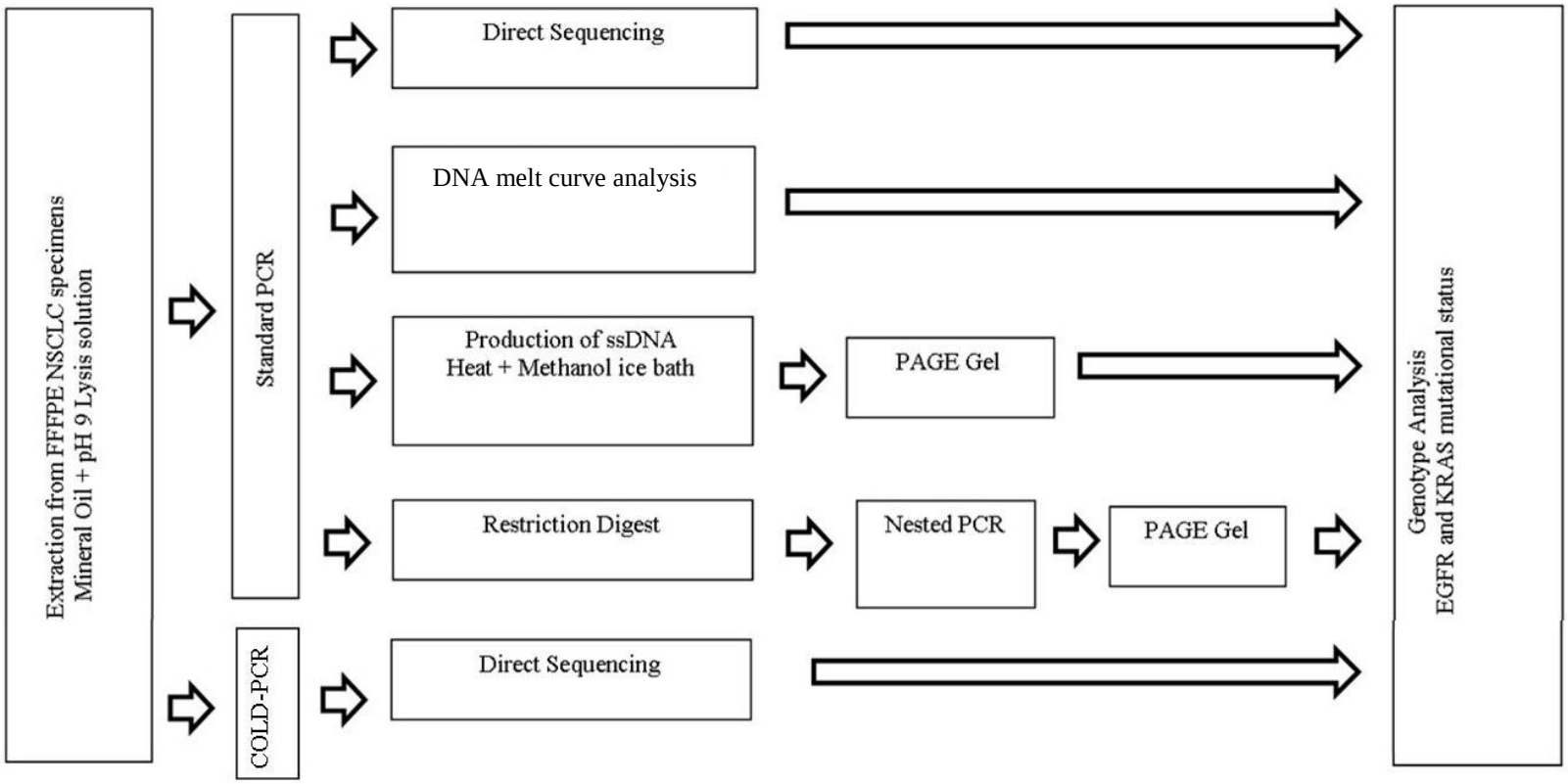

Figure 5 Summary of protocols employed in the detection of *EGFR* exon 19 and exon 21 and *KRAS* exon 2 mutations in NSCLC FFPE specimens

Chapter 2 Materials and Methods

2.1 Sample Collection and Identification

FFPE NSCLC biopsy specimens were selected from the database of archived outpatient specimens at the Waikato Hospital (Hamilton, NZ). The sex, histology and smoking history of these patients was included alongside the specimen reference number. This information was used to select specimens to enrich possibility of specimens containing *EGFR* mutations. Table 2 shows the sex, smoking status of the patient, as well as the histological subtype of the six specimens that were selected for mutation analysis.

Once selected, the seven specimens were analysed by an experienced pathologist, who identified and marked the areas containing tumour and normal tissue within these biopsy specimens. This was done to maximise the collection of tumour DNA over normal tissue to reduce the wild-type DNA background.

Following collection from Waikato hospital, each specimen was kept in a separate labelled petri dish to prevent any cross-contamination of or between specimens.

Table 3 FFPE NSCLC specimens selected from Waikato Hospital Outpatient Database. Patient reference number, sex, smoking status and the histological subtype of their tumours were the only information made available to researchers.

Specimen	Sex	Histology	Smoking Status
AUN6654	Female	Adenocarcinoma	Non-smoker
PSA1523	Female	Adenocarcinoma	Non-smoker
PRF3901	Female	NSCLC	?*
EBT5130	Female	?*	Non-smoker
AWV7836	Female	Adenocarcinoma	Non-smoker
EMV8895	Female	Adenocarcinoma	Non-smoker
DGZ6804	Female	Adenocarcinoma	Non-smoker

***Information was not available.**

2.2 DNA extraction/isolation from FFPE Specimens

A #10 sterile scalpel blade (Swann Morton Ltd) was used to cut 4-6 thin slices from the FFPE specimens. These slices were placed in a labelled 1.7mL microcentrifuge tube (Axygen) containing 300mL mineral oil and 600mL pH 9 SDS-lysis solution (1M Tris pH 9, 50mM EDTA, 1% SDS). The tubes were mixed in a thermomixer (Eppendorf) at 90^{o}C and 900rpm for 20minutes to dissolve the paraffin wax. The thermomixer was then cooled to 56^{o}C before introducing 10µL Proteinase K (0.06g of powder Proteinase K, final conc. 60µg/µL) to each tube and incubating at this temperate at 600rpm for 4 hours. After which, another 10µL Proteinase K was added before incubating at 37^{o}C (Precision) overnight.

The following day, the tube was returned to the thermomixer and heated to 56^{o}C to melt the wax and separate the oil from the specimen, causing the formation of separate phases. The bottom phase containing the DNA was removed using a 1mL transfer pipette (Raylab) and transferred into a new labelled 1.7mL microcentrifuge tube along with an equal volume (approximately 600mL) of phenol-chloroform (Sigma). To ensure the phase containing the oil and paraffin wax remained in the tube, the pipette tip was placed right to the bottom of the tube. The mircocentrifuge tube containing the top oil phase was then discarded.

The new labelled tube was shaken vigorously by hand to allow the phenol-chloroform to mix with the specimen, until an emulsion was formed. After which, the tube was put on the rotator wheel (Global Science) to continue mixing for 20minutes. The tube was then centrifuged (5415R Bench Top Centrifuge, Eppendorf) at 16.1kRCF for 15minutes. The top phase containing the DNA (approximately 400mL) was transferred to new 1.7mL microcentrifuge tube and topped up to 600mL with pH 9 SDS-Lysis solution (approximately 100mL). 100µL 5M NaCl was then added and the specimen was inverted several times to mix. Then, 80mL of pre-heated cetrimonium bromide (CTAB) as added before incubating the specimen at 65^{o}C for 10minutes.

After which, an equal volume of chloroform (Ajax Finechem Pty Ltd) was added, and the tube was placed on the rotator wheel for 20 minutes. Then, the tube was centrifuged at 16.1kRCF for 15minutes. The top phase was then removed and

transferred to a new labelled 1.7mL mircocentrifuge tube. Following this, an equal volume of isopropanol was added and the solution was frozen overnight.

The following day, the tube was centrifuged at 16.1kRCF for 15minutes. After which, the top phase was tipped off and 1mL 70% ethanol was added before returning the tube to the centrifuge for 5minutes at 16.1kRCF. Then, the top phase was once again tipped off and the tube was centrifuged for 5seconds. Finally, a 200µL pipette was used to remove the remaining ethanol and the tube was placed in the fumehood before resuspending in 50µL TE buffer (10mM Tris, 1mM EDTA pH 8).

2.2.1 CTAB Clean up

Those FFPE specimens which had a poor 260/280 ratio (less than 1.4) were subjected to CTAB clean up. Specifically, 470µL TE and 30µL 10% SDS were added to the remaining DNA. This was then placed in the thermomixer and incubated at 65^{o}C and mixed at 900rpm for 10minutes.

After which, 100µL 5M NaCl was added with 80µL of pre-warmed CTAB and further incubated in the thermomixer under the same conditions for 10 minutes. Then, an equal volume of chloroform was added, and the tube was placed on the rotator wheel for 20 minutes. Then, the tube was centrifuged at 16.1kRCF for 15minutes. The top phase was then removed and transferred to a new labelled 1.7mL mircocentrifuge tube. Following this, an equal volume of isopropanol was added and the solution was frozen overnight.

The following day, the tube was centrifuged at 16.1kRCF for 15minutes. After which, the top phase was tipped off and 1mL 70% ethanol was added before returning the tube to the centrifuge for 5minutes at 16.1kRCF. Then, the top phase was once again tipped off and the tube was centrifuged for 5seconds. Finally, a 200µL pipette was used to remove the remaining ethanol and the tube was placed in the fumehood before resuspending in 50µL TE buffer.

2.3 A549 NSCLC adenocarcinoma cell line

The A549 NSCLC adenocarcinoma cell line was developed in 1972 by culturing cancerous lung tissue explanted from a 58 year old Caucasian male. Since, genotype analysis has shown that the A549 cell line is homozygous for the 34G>A (G12S) point mutation in *KRAS* exon 2 (Krypuy, et al., 2006). Indeed, due to the strict mutually exclusive relationship *KRAS* and *EGFR* mutations, these cells are wild-type for activating *EGFR* exon 19 and exon 21 mutations.

Therefore, these cells were used as controls throughout the experimental work as the mutation status of both *KRAS* and *EGFR* genes are known for these cells. Consequently, any specimens showing aberrant activity when compared to the genomic DNA isolated from these cells was considered indicative of the presence of a mutation.

2.4 DNA extraction/isolation from A549 cells

A549 cell cultures were supplied by Dr Ray Cursons. The cell cultures were analysed under the microscope to confirm the presence of A549 cells in the medium. After which, the media was tipped off by inverting the flask, and any remaining media was removed using a 200µL pipette.

Then, 1mL of SDS-lysis solution was added to the flask and gently swirled to ensure the lysis of the cells adhering to the surface of the flask. Following this, the flask was incubated at 37^{o}C for 15minutes. After incubation, 500µL of the solution was placed in a labelled 1.7mL mircocentrifuge tube containing 1mL 5M LiCl (Ajax Finechem Pty Ltd). Then, the tube was mixed until the proteins precipitated and the solution turned cloudy.

Then, approximately 200µL chloroform was added to the mircocentrifuge tube, which was vigorously shaken to form an emulsion. After which, the mircocentrifuge tube was placed on the rotator wheel for 20minutes. Following this, the microcentrifuge tube was transferred to the centrifuge and centrifuged for 10minutes at 16.1kRCF.

Then, using a 1mL transfer pipette the aqueous top layer containing the genomic DNA was removed and transferred to a new 1.7mL microcentrifuge tube.

Following this, 500µL isopropanol was added to the tube and the tube was centrifuged for 20minutes at 16.1kRCF.

Following centrifugation, the supernatant was removed and discarded using a 1mL transfer pipette, leaving the pellet containing the genomic DNA in the tube. Then, the pellet was re-suspended in 1mL 70% ethanol and centrifuged for 5minutes at 16.1kRCF. After which, the ethanol was tipped off and the tube was returned to the centrifuge for 15seconds. Then, using a 200µL pipette the excess ethanol was removed before placing the mircocentrifuge tube in the fumehood to evaporate the remaining ethanol. Following this, the pellet was re-suspended in 50µL TE buffer.

2.5 DNA concentration and Purity

The DNA concentration and purity of each specimen was determined by using 2µL for analysis using a ND-1000 spectrophotometer (Nanodrop). The concentration of nucleic acids was measured by absorbance at 260nm and given in ng/µl, while the ratio of 260/280nm indicates the purity of the sample. A ratio of 1.8-2.0 was ideal, yet is only informative in terms of contamination and does not provide information on any DNA degradation that can occur as a result of fixing the specimen in formalin.

The DNA extraction procedure from most specimens was done in duplicate (except 19A8 as there was not enough tissue). After the specimens were analysed in the Nanodrop, the concentration of genomic DNA was diluted to approximately 100ng/µL for use in PCR.

2.6 Amplification of *EGFR* Exon 19 and 21 and *KRAS* Exon 2

Based on previous publications (Krypuy, et al., 2006; Nomoto, et al., 2006), primers pairs specific for *EGFR* exon 19 and 21 and *KRAS* exon 2 were synthesised (IDT).

The primers were supplied at concentrations between 24.0 - 40.5nmol and were diluted to 200pmol with TE buffer. For each exon, working primer solutions

containing both the forward and reverse primers were made to 20pmol/µL concentration.

Table 4 Standard PCR Primers for genotyping using *EGFR* exon 19 and 21 and *KRAS* exon 2. The base-pair (bp) length of each target region is given.

Exon	Forward (5' – 3')	Reverse (5' – 3')	Length (bp)
EGFR Exon 19	AAAATTCCCGTCGCTATC	AAGCAGAAACTCACATCG	276
EGFR Exon 21	AGATCACAGATTTTGGGC	ATTCTTTCTCTTCCGCAC	224
KRAS Exon 2	TCATTATTTTTATTATAAGGCCTGCTGAA	CAAAGACTGGTCCTGCACCAGTA	189

A standard PCR protocol was implemented for the amplification of *EGFR* exon 21 and 19 and *KRAS* exon 2 using the respective set of primers as per Table 4.

2.6.1 Optimisation of PCR

Optimisation of amplification during PCR is important to address the myriad of issues when dealing with FFPE specimens, given the quality and quantity of the genomic DNA available for these assays is not ideal. Indeed, the optimisation of PCR can significantly increase the sensitivity of downstream DNA analysis. As such, the conditions for the PCR thermocycling were undertaken with different magnesium concentrations as well as with DNA *Taq* polymerase enzymes possessing different fidelities.

Table 5 Standard PCR protocol for amplification of *EGFR* exon 21 and 19 and *KRAS* exon 2

Step		Temp (^{o}C)	Time
1	Initial denaturation and enzyme activation	95	15minutes OR 5 minutes
2	Denaturation	95	30seconds
3	Primer annealing	55	30seconds
4	Extension	72	20seconds
5	Repeat Steps 2 – 4 for additional 39 cycles		

*** Length of activation step different for specific DNA polymerase enzymes. HOTFIRE *Taq* polymerase required a 15minute initial activation while FIREPOL *Taq* polymerase only required a 5minute initial activation.**

As such, PCR was performed as per Table 5, with 0.4µM of each primer, 200µM of each dNTP (A, G, C and T, Solis Biodyne), either 1.5mM or 4.5mM $MgCl_2$ and 1.25U/50µL of FIREPOL or HOTFIRE *Taq* Polymerase enzyme (Solis Biodyne) in a final reaction volume of 50µL. The PCR protocol was performed on a Thermal cycler (MJ Research).

2.6.2 Agarose Gel Electrophoresis

The resultant PCR products from Standard PCR were electrophoresed and visualised on an Owl gel electrophoresis system to determine amplification of the target exons.

0.66g agarose (SeaKem®) was added to 33mL SB (56g Boric acid, 10g NaOH) and heated with intermittent stirring until the agarose had dissolved in the solution and no crystals could be seen. After which, the solution was cooled under running cold water and 1.25µL ethidium bromide (usb) (3µL/30ml of 10mg/mL stock solution) was added to the solution and mixed well. Then, the solution was poured into a gel castor and two combs were inserted and secured in place. The gel was left to set for at least 30minutes.

10µL of the amplified product from each PCR reaction was mixed with 2µL loading dye (0.05% bromophenol blue (Ajax Chemicals), 0.05% Xylene cyanol (Sigma), 6% glycerol (Ajax Chemicals)) and loaded into a well of the gel. A 100bp ladder (Solis BioDyne) was also loaded into one of the wells for comparison of PCR product length. The products were then electrophoresed for

30mins at 110V and 30mA. Product bands were visualised using a ultra-violet light transluminator (TFX-35M, Life technologies), photographed (COHU High Performance Camera) and printed.

2.6.3 Purification of PCR products for sequencing

Following visualisation on the agarose gel, PCR products were purified by precipitation with polyethylene glycol (PEG). As such, approximately 40µL of the PCR product was transferred to a 1.7mL microcentrifuge tube and an equal volume of PEG solution (20% polyethylene glycol 800 in 2.5mM NaCl) was added, and the tube was inverted several times to ensure the solution was mixed well. Then, the tube was incubated at room temperature for 15minutes.

After which, the microcentrifuge tube was centrifuged at 16.1kRCF for 15minutes to cause the DNA to sediment. Then, the supernatant was removed leaving the pellet in the tube, and 1mL 100% ethanol was added before centrifuging at 16.1kRCF for 5minutes. Following this, the ethanol was removed and 1mL 70% ethanol was added before inverting the tube several times and centrifuging at 16.1kRCF for another 5minutes. Then, the ethanol was tipped off and the tube was centrifuged at 16.1kRCF for 10seconds and the remaining ethanol was removed using a 200µL pipette.

Finally, the pellet was left to dry in the fumehood, before resuspending in 10µL Milli-Q H_2O (Barnstead). The tube was then vortexed to ensure the solubilisation of the DNA and the concentration of DNA was measured using the ND-1000 spectrophotometer. The solutions were then sent to the University of Waikato Sequencing Facility (Hamilton, NZ), using Applied Biosystems 3130xl Genetic Analyser and the forward and reverse primers for the respective exons.

2.7 COLD-PCR

COLD-PCR takes advantage of the substantial drop in amplification of homo-duplexes when the denaturation temperature is dropped below the critical denaturation temperature (T_c). This enables the selective amplification heteroduplexes of mutant and wild-type sequences and enriches the original minority mutant sequences (Luthra & Zuo, 2009). This protocol was used in attempt to increase the sensitivity of both direct sequencing and high resolution melting analysis.

2.7.1 Experimental Determination of T_m and T_c

To experimentally determine the T_m and T_c of the products produced for COLD-PCR using the respective primers for each exon (Table 4), the optimal conditions as determined in 2.6.1 were set up, with some modifications for Real-time PCR. Accordingly, less primer and *Taq* DNA polymerase enzyme were added and the fluorescent dye, Syto 82 (Invitogen) was introduced.

As such, Real-time PCR was performed with 0.25µM of each primer, 200µM dNTPs, 4.5mM $MgCl_2$ and 0.6U of HOTFIRE *Taq* Polymerase enzyme and 0.025mM Syto82 in a final reaction volume of 20µL. The PCR was performed on the Real Time Cycler (RotoGene^{TM}6000, Corbett Technologies) and melt curves where produced using the associated software. These melt curves gave both the melting activity and the T_m of the amplicons from which the T_c could be determined using the following equation:

$$Tc(℃) = \mathrm{Tm}(℃) - 1℃$$

After the T_c of the *EGFR* exon 21 and 19 and *KRAS* exon 2 amplicons were determined, thermocycling was undertaken with the denaturation temperature set to the respective T_c.

2.7.2 COLD-PCR amplification of EGFR Exon 19 and 21 and KRAS Exon 2

Reaction conditions for the COLD-PCR amplification of *EGFR* exon 19 and 21 and *KRAS* exon 2 were similar to the conditions outlined in 2.6.1 with slight modifications. Specifically, 15 cycles 'normal cycles' where carried out before dropping the denaturation temperature to T_c for an additional 40 cycles, given that beginning with normal cycles is essential to increase the template concentration for thermocycling.

Table 6 COLD-PCR Protocol for amplification of *EGFR* exon 21 and 19 and *KRAS* exon 2

Step		Temp (^oC)	Time
1	Initial denaturation and enzyme activation	95	15minutes
2	Denaturation	95	15seconds
3	Primer annealing	55	10seconds
4	Extension	72	15seconds
5	Repeat Steps 2 – 4 for additional 14 cycles		
6	Denaturation	T_c*	15seconds
7	Primer annealing	55	10seconds
8	Extension	72	15seconds
9	Repeat Steps 6 – 8 for 39 more cycles		

***T_c experimentally determined as per 2.7.1 for each respective exon**

As such, PCR was performed on a Thermal Cycler as per Table 6, with 0.4µM of each primer, 200µM dNTPs, 4.5mM $MgCl_2$ and 1.25U/50µL of *Taq* Polymerase enzyme in a final reaction volume of 50µL.

Once the COLD-PCR protocol was carried out, 10µL of each reaction was loaded into a 2% (w/v) agarose gel supplemented with ethidum bromide and electrophesed as previously described in 2.6.2. Following this, the COLD-PCR products were purified and sent to the University of Waikato DNA Sequencing Facility, as per 2.6.3.

2.8 Restriction Digest-mediated mutant enrichment PCR

Restriction digest mutant enrichment PCR is a two-step PCR with intermittent restriction digestion to eliminate wild-type alleles selectively, thus enriching the mutated alleles. Restriction enzymes have recognition sites at which they selective cut or 'digest' a DNA sequence. As such, restriction enzymes, which have a recognition site within the wild-type gene, can be used to selectively digest these sequences, increasing the proportion of mutant sequences that can be used for templates during amplification in PCR.

Indeed, restriction enzymes with recognition sites present exclusively in the wild-type sequences for *EGFR* exon 19 and 21 were selected (Table 6). The restriction enzyme *Mse*I employed for *EGFR* exon 19 has a recognition site (TTAA). This sequence is the first letter of codon 747 (i.e. the first T) to the first letter of codon 748 (i.e. the last A), which is not present in the LREA deletion mutant. While, *Msc*I employed for *EGFR* exon 21 as it has a recognition site (TGGCCA) sequence which is not present in the L858R mutant due to the base substitution of T to G at first base of recognition site.

Table 7 Restriction enzymes selected for mutant enrichment PCR of *EGFR* exon 19 and exon 21. Both restriction enzymes have recognition sites that are not present in the mutant.

Exon	Restriction Enzyme	Recognition Site	Reason why not present in mutant sequence
EGFR Exon 19	*Mse*I	TTAA	First letter of codon 747 to first letter of codon 748 – absent in deletion mutant
EGFR Exon 21	*Msc*I	TGGCCA	GGGCCA in L858R mutant

2.8.1 Restriction digest-mediated mutant enrichment PCR amplification of *EGFR* Exon 19 and 21

The method employed was a combination of those previously described by Asano et al (2006) and Hlinkova et al (2011). Specifically, the first round of PCR amplification was done in a 20μL volume with the primers in Table 10. Importantly, the second round of PCR was done with a second set of primers

inside the first round primers. Reaction conditions for the amplification of *EGFR* exon 19 and 21 were similar to the conditions outlined in 2.6.1 with slight alterations.

Table 8 Restriction Digest-mediated mutant enrichmnent PCR protocol for amplification of *EGFR* exon 21 and 19

Step		Temp (^{o}C)	Time
1	Initial denaturation and enzyme activation	95	15minutes
2	Denaturation	95	30seconds
3	Primer annealing	53	30seconds
4	Extension	72	20seconds
5	Repeat Steps 2 – 4 for additional 39 cycles		
6	Final extension	72	10 minutes

As such, Restriction digest-mediated mutant enrichment PCR was performed on a Thermal Cycler as per Table 8, with 0.4 µM of each primer, 200µM dNTPs, 4.5mM $MgCl_2$ and 1U/50µL of HOTFIRE *Taq* Polymerase enzyme in a final reaction volume of 20µL.

2.8.2 Restriction enzyme digestion

The restriction digest of the first PCR products for *EGFR* exon 21 was done using 2U of *Msc*I (BioLabs). Specifically, a 5µL aliquot of the restriction buffer as per Table 9 was added to the 20µL first round PCR reaction.

The intermittent restriction digest of the first PCR products for *EGFR* exon 19 was carried out in a similar way using 2U of *Mse1* (BioLabs), 0.25µL BSA and only 1.85µL Milli-Q H_2O was added to the restriction digest buffer solution as per Table 9.

The tube was incubated at 37^{o}C for 4 hours. Then, the restriction enzyme was deactivated by increasing the temperature to 65^{o}C for 20minutes.

Table 9 Restriction digestion incubation conditions for *EGFR* Exon 21 and Exon 19

Reagents	Volume	
	Exon 21	**Exon 19**
10x Buffer	2.5µL	2.5µL
Milli-Q H_2O	2.1µL	1.85µL
PCR Product	20µL	20µL
Enzyme*	0.4µL	0.4µL
BSA	-	0.1µL

***Restriction enzyme was specific to each exon**

2.8.3 Nested PCR

Nested PCR achieves greater PCR sensitivity and specificity through the use of a second round of PCR. As such, an aliquot of the PCR product is used as the template DNA for a second round of PCR with primers designed to amplify a smaller DNA region within the primary amplicon i.e. the second primers are 'nested'.

The second round of PCR was performed with 2µL of the digestion product from 2.8.2 as the template under the same conditions as the first round of PCR, except for two modifications. Specifically, the use of the nested primers (Table 9) and thermocycling was only carried out for 35 cycles. The PCR products from the second amplification were then analysed on a 13% polyacrylamide gel electrophoresis (PAGE) gel with silver staining as described in 2.9.2 and 2.9.4.

Specifically, 10µL of the PCR products was mixed with 2µL of loading dye and loaded into a well of the gel. A 100bp ladder (BioDyne) was also loaded into the gel for restriction fragment length polymorphism analysis.

Table 10 Nested Primers for Restriction digest-mediated mutant enrichment PCR

Exon	Forward (5' – 3')	Reverse (5' – 3')
Round 1		
Exon 19	GCACCATCTCACAATTGCCAGTTA	CCTGAGGTTCAGAGCCATGGA
Exon 21	CCATGATGATCTGTCCCTCACA	TCCCTGGTGTCAGGAAAATGCT
Round 2		
Exon 19	ATCCCAGAAGGTGAGAAAGATAAATTC	GAGGTTCAGAGCCATGGACCC
Exon 21	CGCAGCATGTCAAGATCACAGAT	AGGAAAATGCTGGCTGACCTAAAG
Product lengths – exon 19 round 1 = 176bp, exon 19 round 2 = 114bp – exon 21 round 1 = 194bp, exon 21 round 2 = 122bp		

2.9 Single-stranded Conformation Polymorphism

PCR-SSCP is a well-established method used for mutation screening and detection in biological specimens, where mutations are often present in at a low frequency. This method relies on the fact that a mutation along a sequence will result in that sequence adopting a different secondary conformation when heated and cooled rapidly.

2.9.1 Production of ssDNA

PCR products for *EGFR* exon 19 and 21and *KRAS* exon 2 were produced as per 2.6.1 for the use in SSCP. After amplification, 5µL the PCR products were diluted with 15µL of stop buffer (95% formamide 0.05% bromophenol blue, 0.05% xylene cyanol) and denatured at 95^{o}C for 15minutes, and then chilled for 10 minutes in a methanol/ice-bath (-20^{o}C).

2.9.2 PAGE gels

Polyacrylamide gels were set up using a Hoefer mighty small SE250/SE260 gel rig. The glass plates and spacers were cleaned with detergent and tap water, wiped dry with paper towels, and then cleaned again with 70% ethanol to ensure surfaces were free of any polymerised acrylamide and other debris. Prepared plates were then carefully placed in a Hoefer gel caster. Gels were

prepared according to the appropriate composition shown in Table 9. Ammonium persulfate (APS) and TEMED (N, N,N,N',N'-tetramethyl-ethylenediamine) were not added until immediately before casting the gel. The gel was prepared in a 50 ml falcon tube (CELLSTAR®) then poured immediately into the gel caster using a 1mL transfer pipette. When the gel caster was full, a 10 tooth comb was put in place. Extra solution was added around the comb to ensure the formation of even and deep wells. The gels were observed for a few minutes to ensure no leakage. The gel was left to set and if was not being used that day were wrapped in gladwrap and stored at 4°C for no longer than 1 week.

Table 11 Composition of 13% PAGE gels

Reagent	Volume
49:1 Acrylamide:bis (40%)	3mL
5xTBE	2mL
50% glycerol	1mL
H2O	3mL
TEMED	10µL
10 % APS (w/v)	100µL
Sufficient to prepare one 10 x 8 gel.	

2.9.3 Electrophoresis of ssDNA

When required, gels were unwrapped and clamped to a Hoefer 'Mighty Small II' (SE250/SE260) gel electrophoresis unit with a cooling plate. The buffer chamber was filled to the appropriate level with 1X TBE (0.45 M Tris, 0.45 M Boric Acid, 10mM EDTA) buffer. The comb was removed and wells washed to remove any unpolymerised acrylamide from the wells with 1XTBE buffer using a 200µL pipette.

Then, 10µL aliquots of the prepared samples from 2.9.1 were loaded onto the polyacrylamide gel with a 10µl pipette. The running conditions were 200 V (~30mA) for 3-5 hours at room temperature.

2.9.4 Silver Staining of acrylamide gels

After electrophoresis, the gels were silver stained according to Dr Ray Curson's protocol (2005). Accordingly, the gels were removed from their gel plates and washed in fixative solution (10% ethanol 5% acetic acid) for 3 minutes. Then, the fixative solution was removed using the suction pump before soaking the gel in 1% HNO_3 for 3 minutes. Once again, the suction pump was used to remove the solution and the gel was rinsed in Milli-Q H_2O for 30 seconds. After the Milli-Q H_2O was removed with the suction pump, the gel was soaked in 0.1% $AgNO_3$, 150µL HCOOH/100mL for 10 minutes. This was removed via the suction pump before rinsing the gel once again in Milli-Q H_2O for 30 seconds and this was removed. The gels were then allowed to develop in the developer solution (3% Na_2CO_3, 150µL HCOOH, 100µL 10% Nathio) until stained bands were visiable. The development was stopped by addition of fixative solution (10% acetic acid). The gels were then washed in Milli-Q H_2O and dried before photographed over a light box and sealed in a plastic bag.

2.10 DNA melt curve analysis

DNA melt curve analysis has recently been introduced as a rapid, sensitive an inexpensive method for the screening of mutations in a substantial wild-type background. As such, it has huge potential for detection of DNA sequence changes such as mutations and insertions and deletions (Do et al 2008).

2.10.1 Amplification of *EGFR* Exon 19 and 21 and *KRAS* exon 2 with Real time PCR

Real-time PCR for melt curve analysis was performed on the Rotor-Gene 6000 (Corbett Research) in the presence of the fluorescent DNA intercalating dye, Syto 82 (Invitrogen) and primers previously described in Table 4.

Table 12 Real-time thermocycling parameters for amplification of *EGFR* exon 21 and 19 and *KRAS* exon 2

Step		Temp (^{o}C)	Time
1	Initial denaturation and enzyme activation	95	15minutes
2	Denaturation	95	30seconds
3	Primer annealing	55	30seconds
4	Extension	72	20seconds
5	Hold	80	10seconds
6	Repeat Steps 2 – 5 for additional 39 cycles		

As such, Real-time PCR was performed as per Table 12, with 0.25μM of each primer, 200μM dNTPs, 4.5mM $MgCl_2$ and 0.625U of HOTFIRE *Taq* Polymerase enzyme and 0.025mM Syto82 in a final reaction volume of 20μL.

2.10.2 Production of Melt Curves

Following PCR, products from 2.10.1 were heated from 60°C to 95°C, rising by at 0.1-0.2^{o}C/sec and the change in fluorescence, as measured by release of Syto82, was recorded with the accompanying software.

2.10.3 Melt Curve Analysis

The melt curves generated from the FFPE specimens in 2.10.3 were compared to the melt curves generated using the same protocol for the A549 cell line. Samples with a skewed or left-shifted curve from those of control samples are judged to have a mutation.

2.11 Analysis of Sequencing Data

The sequences obtained from the University of Waikato DNA Sequencing facility were checked against sequences stored at online database 'Genebank' by a BLAST analysis. Mutations were identified by comparison between the online reference sequences for *EGFR* (NG_007726.1) and *KRAS* (NG_007524.1) genes on this database and the sequences obtained from the specimens. As such, any misalignments were further analysed to elucidate the exact position and consequently establish if the mutation was of clinical significance or just an experimental artefact.

Chapter 3 Results

3.1 Isolation of genomic DNA

Genomic DNA was extracted and isolated from the seven FFPE NSCLC specimens using a method adapted from Lin et al (2009). Importantly, our method utilised mineral oil and a pH 9 SDS-lysis solution for the efficient removal of paraffin wax, contributing to the extraction of a high yield of quality genomic DNA from the FFPE specimens.

Table 13 Nanodrop results from DNA extraction and isolation carried out on seven FFPE NSCLC specimens and the A549 NSCLC adenocarcinoma cell line. The physical characterisation of the specimens was also recorded.

Specimen	Concentration (ng/µL)	260/280 Absorbance	Physical Characterisation of FFPE specimen
24A6	319.7	1.95	Mostly tumour tissue
54A4	116.7	1.89	Mostly tumour tissue
98F7	743.4	1.82	All tumour tissue
77B4	634.7	1.84	Mostly tumour tissue
34F3	847.9	1.89	Mostly normal tissue
4674	672.3	1.85	Mostly tumour tissue
19A8	64.8	1.83	Mostly tumour tissue – very small specimen
A549 cells*	953.9	1.87	N/A

***A549 NSCLC adenocarcinoma cell line is not a FFPE specimen**

Table 13 shows the Nanodrop results from one round of DNA extraction (all specimens were done in duplicate except 19A8 due to a lack of tissue). The concentration of DNA isolated from these specimens was surprisingly high and the 260/280 absorbance reading was close to 1.8, indicating a lack of contaminants within the DNA sample. However, two specimens, 77B4 and 98F7 had a low 260/280 reading which indicated the extraction protocol was not able to

sufficiently remove other nucleic acids and/or there was some remaining reagents left in the DNA pellet. As a consequence, these specimens where subjected CTAB clean-up as per 2.3.1 to remove any contamination and the DNA concentration and 260/280 absorbance were recalculated (as shown in Table 13).

Moreover, while most FFPE specimens yielded a significant amount of DNA specimen 19A8, which was characterised as having a very small amount of tissue (typical of fine needle aspirations from NSCLC patients) yielded significantly less genomic DNA. Given that approximately 100ng is required as a DNA template for PCR, this low yield of DNA is likely to limit the downstream DNA analysis for this specimen.

3.2 Optimisation of Standard PCR

PCR thermocycling was carried out with differing time periods, temperatures and concentrations of reagents to find the optimum conditions to maximise amplification in attempt to increase both the sensitivity and robust nature of the downstream mutation detection protocols. Specifically, we found that using both a PCR buffer with high magnesium concentration and high fidelity *Taq* DNA polymerase enzyme were important to ensure efficient amplification from the FFPE specimens. Interestingly, these optimised conditions were not required for amplification of DNA isolated from the A549 NSCLC adenocarcinoma cell line, further highlighting the importance of optimising PCR conditions when amplifying a degraded DNA template.

3.2.1 Magnesium Concentration

To establish the effect of magnesium concentration on the efficiency of PCR amplification from FFPE specimens, the standard PCR protocol as described in 2.6.1 was carried out with a PCR buffer containing a standard $MgCl_2$ concentration of 1.5mM and with a PCR buffer with a higher concentration of 4.5mM $MgCl_2$. Figure 6 and Figure 7 show the results from agarose gel electrophoresis conformation of these PCR conditions.

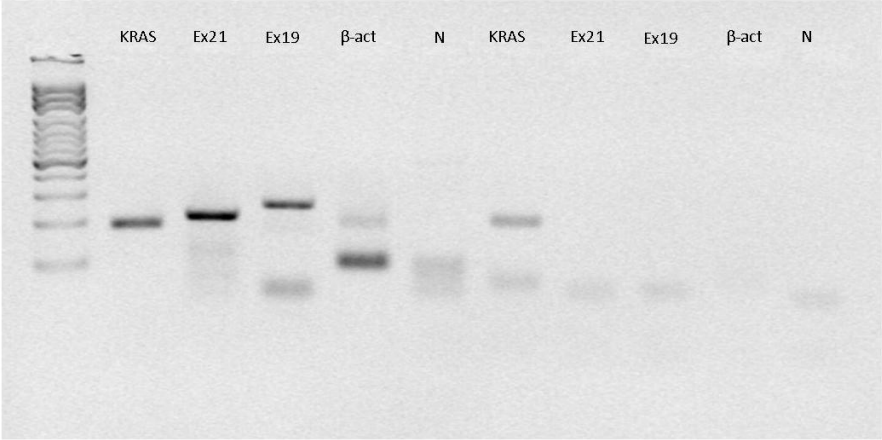

Figure 6 Standard PCR amplification of DNA isolated from A549 NSCLC adenocarcinoma cell line (Lanes 2-6) and FFPE specimen (Lanes 7-11) using a PCR buffer with a 1.5mM magnesium concentration. *KRAS* exon 2, Ex21 = *EGFR* exon 21, Ex19 = *EGFR* exon 19, ß-act = *ß-actin* (housekeeping gene). Lane 1 = 100bp ladder.

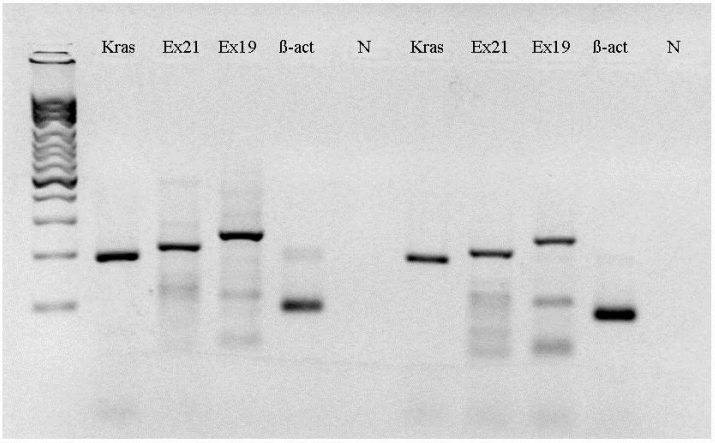

Figure 7 Standard PCR amplification of DNA isolated from FFPE specimens (Lanes 2-6) and A549 cell line (Lanes 7-11) using a PCR buffer with a 4.5mM magnesium concentration. KRAS = *KRAS* exon 2, Ex21 = *EGFR* exon 21, Ex19 = *EGFR* exon 19, ß-act = *ß-actin* (housekeeping gene) and N = negative control. Lane 1 = 100bp ladder.

3.2.1.1 Conclusions

As shown in Figure 6, the PCR buffer possessing a 1.5mM $MgCl_2$ concentration was unable to produce amplification products for *EGFR* exon 21 and 19, *KRAS* exon 2 and *ß-actin* (housekeeping gene) from the FFPE specimens, yet was able to amplify the genomic DNA isolated from the A549 NSCLC adenocarcinoma cell line. In contrast, the PCR buffer possessing a 4.5mM $MgCl_2$ concentration was able to amplify products from the genomic DNA isolated from both the A549 cells and the FFPE specimens (Figure 7). Consequently, all subsequent experiments were carried out with a PCR buffer with a 4.5mM $MgCl_2$ concentration.

3.2.2 High fidelity *Taq* DNA polymerase enzyme

The requirement of a high fidelity *Taq* DNA polymerase enzyme was determined by carrying out the standard PCR protocol as described in 2.6.1 using FIREPOL *Taq* DNA polymerase and HOTFIRE *Taq* DNA polymerase. HOTFIRE *Taq* DNA Polymerase is a chemically modified FIREPOL DNA polymerase. At ambient temperatures it is inactive, having no polymerisation activity as the HOT FIREPol DNA polymerase is activated by a 15 min incubation step at 95°C. This prevents extension of non-specifically annealed primers and primer-dimers formed at low temperatures during PCR. The enzyme has 5'→3' polymerization-dependent exonuclease replacement activity but lacks 3'→ 5' exonuclease activity (Solis Biodyne Datasheet).

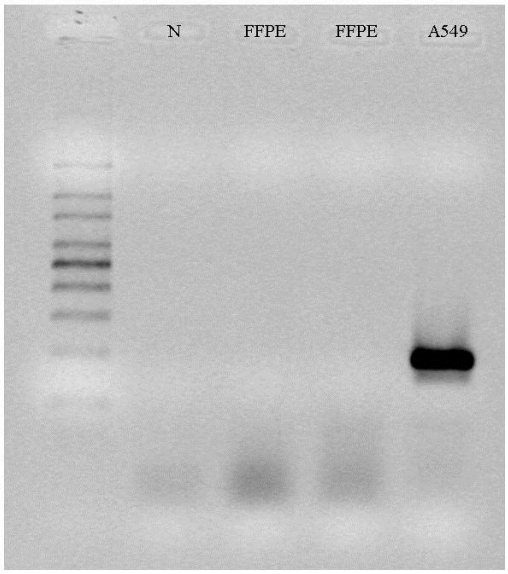

Figure 8 Standard PCR amplification of *KRAS* exon 2 from genomic DNA isolated from FFPE specimens (Lanes 3 and 4) and the A549 NSCLC adenocarcinoma cell line (Lane 5) using FIREPOL *Taq* DNA polymerase enzyme. N = negative control, no DNA. Lane 1 = 100bp ladder.

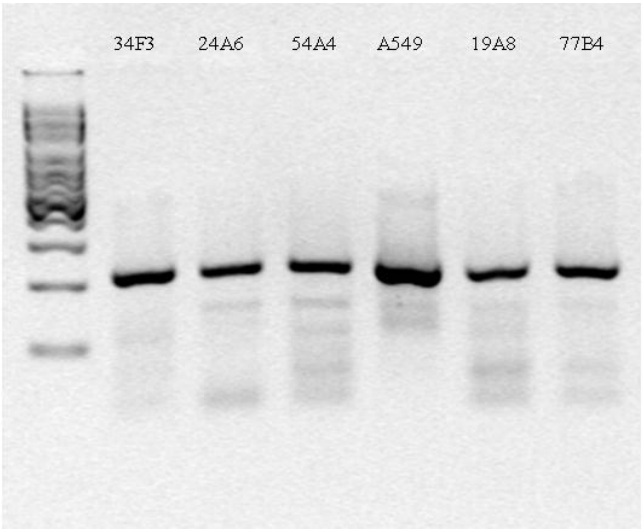

Figure 9 Standard PCR amplification of *KRAS* exon 2 from FFPE specimens (Lanes 2 -4 and 6-7) and A549 adenocarcinoma cell line (Lane 5) using HOTFIRE *Taq* DNA polymerase enzyme. Lane 1 = 100bp ladder.

3.2.2.1 Conclusions

Figures 8 and 9 depict the agarose gel electrophoresis conformation from the standard PCR amplification of *KRAS* exon 2 from the FFPE specimens and A549 cell line, using these two *Taq* DNA polymerases. As seen in Figure 8, FIREPOL *Taq* polymerase was only able to produce amplification products from the genomic DNA isolated A549 cells, while HOTFIRE *Taq* polymerase was able to produce amplification products from the DNA isolated from both the A549 cells and the FFPE specimens (Figure 9). Consequently, all the subsequent experiments were carried out using HOTFIRE *Taq* DNA polymerase.

3.2.3 Confirmation by DNA Sequencing

Following the elucidation of the optimum thermocycling conditions, the PCR products for *EGFR* exon 21 and 19 and *KRAS* exon 2 from each of the six NSCLC FFPE specimens were purified and sent to the Waikato DNA Sequencing facility for direct sequencing to determine if standard PCR followed by direct sequencing could detect *EGFR* and *KRAS* mutations in the NSCLC FFPE specimens. Examples of the electropherograms received for each exon are presented in Figures 10-12.

Both forward and reverse DNA sequences were obtained and checked against references stored at online database 'Genebank' by a BLAST analysis. Mutations were identified by comparison between the online reference sequences for *EGFR* and *KRAS* genes on this database and the sequences obtained from each FFPE specimen. As such, any misalignments were further analysed to elucidate the exact position and consequently establish if the mutation was clinically significant (present in the 'mutation hotspots' of each gene) or an experimental artefact.

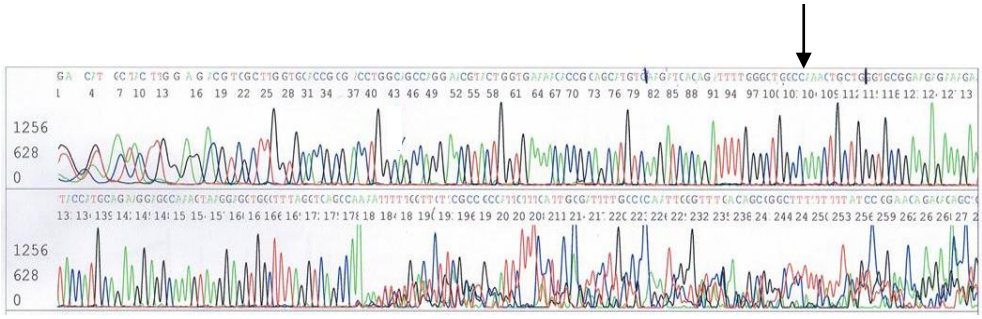

Figure 10 Electropherogram produced from 34F3 NSCLC FFPE specimen for *EGFR* exon 21. The black arrow indicates where the L858R mutation would be present in the exon. In the wild-type sequence (as seen here) a lysine (CTG) is encoded by codon 858 however in the mutated sequence this changes to an arginine (CCG) at this position. This electropherogram is representative of those received for *EGFR* exon 21 as all specimens except 19A8 (no sequence data for standard PCR protocol) showed wild-type sequences in the sequencing data produced from both standard PCR and COLD-PCR.

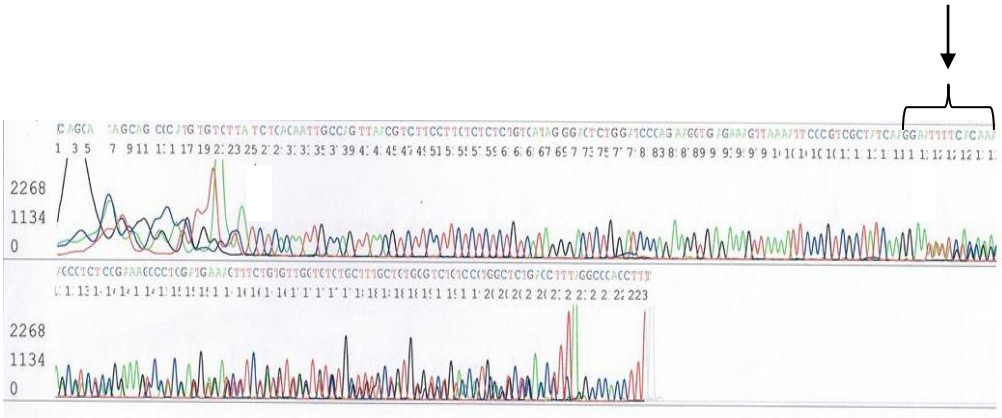

Figure 11 Electropherogram produced from 54A4 NSCLC FFPE specimen for *EGFR* exon 19. The black arrow indicates were the LREA deletion would be present in the exon. In the wild-type sequence this conserved region is present (as shown here), however in the mutated sequence there is 12-`15bp deletion. This electropherogram is representative of those received for EGFR exon 19 as most of the NSCLC FFPE specimens showed a wild-type sequence.

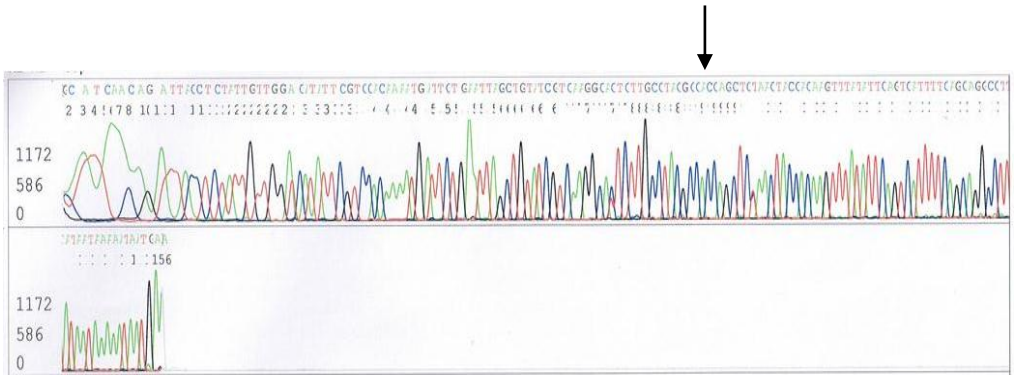

Figure 12 Electropherogram produced from 98F7 NSCLC FFPE specimen for *KRAS* exon 2. The black arrow indicates were the single point mutation would be present in the exon. In the wild-type sequence a glycine (GGT – ACC as reverse sequence shown here) is encoded at codon 12, however in the mutated sequence this changes to a different amino acid (outlined in Table 1). This electropherogram is representative of those received for *KRAS* exon 2 as all specimens showed wild-type sequences in the sequencing data produced from both standard PCR and COLD-PCR.

Indeed, of the sequences obtained from the Waikato Sequencing Facility, none showed any misalignment indicative of the LREA deletion (*EGFR* exon 19), L858R (*EGFR* exon 21) or a codon 12 or 13 single point mutation (*KRAS* exon 2). Examples of the BLAST alignments along with the sequence data (Table 23) for the FFPE specimens are in Appendix III. From this data, the genotypes for the FFPE specimens, as according to standard PCR and direct sequencing were established and presented in Table 14.

Table 14 Genotypes for the six NSCLC FFPE specimens as determined by standard PCR and direct sequencing

Specimen	EGFR Exon 21	EGFR Exon 19	KRAS Exon 2
24A6	WT	WT	WT
54A4	WT	WT	WT
98F7	WT	WT	WT
77B4	WT	?	WT
34F3	WT	WT	WT
19A8	?	?	WT

WT = wild-type ? = unable to genotype

Interestingly, the electropherograms received for the forward and reverse sequencing reactions from 77B4 specimen (that we were unable to obtain a 'clean' sequencing result for *EGFR* exon 19), both stopped at the codons that directly correspond to the mutation hotspot in *EGFR* exon 19 were there the LREA deletion is present. Specifically, the forward sequencing data stopped at codon 746 and the reverse sequencing data stopped at codon 751. This is suggestive of the presence of mutation in this specimen, yet due to the sensitivity limit of direct sequencing, this protocol was unable to detect this deletion in the significant wild-type background.

Further, in attempt to obtain 'clean' electropherogram data and establish a genotype for 77B4 *EGFR* exon 19 and 19A8 *EGFR* exon 21 and exon 19, the PCR products for these specimens were inserted into a plasmid and cloned in DH5α *E. coli* cells (method not shown). However, the sequencing data received from the PCR products of the cloned plasmids was also poor in that

there was a substantial background in the electropherogram, from which a 'clean' sequence could not be obtained and consequently, the genotypes could not be established.

3.2.4 Conclusions

While this protocol is considered the 'gold standard' mutation detection technique and is relatively straightforward and a simple to implement, the standard PCR and direct sequencing protocol was unable to detect any mutations in *EGFR* and *KRAS* genes in the genomic DNA isolated from the FFPE NSCLC specimens. Given that the PCR conditions, while optimised, did not address the sensitivity limit of direct sequencing, further PCR-based protocols must be established to enrich for mutations and/or have greater sensitivity than the reported 20% of direct sequencing.

Indeed, these results are consistent with previous reports (Hlinkova, Babal, Berzinec, Majer, & Ilencikova, 2011b; Krypuy, et al., 2006; Milbury, et al., 2009; Miyamae et al., 2010), in that standard PCR followed by direct sequencing is unable to detect low level mutations in FFPE specimens, due to the poor sensitivity limit of direct sequencing and the inability of PCR to differentially amplify mutant sequences. Importantly, while my results report that all the NSCLC FFPE specimens for which genotype data could be obtained, were wild-type for both *EGFR* and *KRAS*, it is also well established that the use of this protocol does result in a number of false negatives (Takano et al., 2007).

3.3 COLD-PCR

It has been reported that COLD-PCR can increase the sensitivity of PCR-based mutation detection assays by up to 100-fold. Indeed, by changing a single parameter of PCR (the denaturation temperature), this method is relatively easy to implement and is cost-effective in terms of reagents and equipment required.

3.3.1 Experimentally derived T_c

The crucial step in this protocol is the establishment of the T_c for each of the amplicons, which is used as the lower denaturation temperature during thermocycling. As such, the NSCLC adenocarcinoma A549 cell line was used to experimentally determine the critical denaturation temperature (T_c) of the amplicons produced for *EGFR* exons 19 and exon 21 and *KRAS* exon 2. Importantly, A549 cells are wild-type for *EGFR* activating mutations and homozygous for *KRAS* codon 12 mutations. As such, these cells provided an ideal way to experimentally determine the lowest denaturation temperature at which the homoduplexes would separate.

The results of the experimental determination of T_c from 2.7.1 are presented in Table 15. The T_m was established as the melting temperature of the homoduplexes for each of the three exons and the T_c was calculated from this using the previously described calculation in 2.7.1.

Additionally, while our experimentally determined T_c for the *KRAS* amplicon was found to be 81.6^{o}C, we used 81.0^{o}C as we still had good PCR efficiency at this temperature. Indeed, Pritchard et al (2009) found that having T_c set to 81.0^{o}C increased the proportion of *KRAS* mutant alleles amplified in their serial dilution experiments.

Table 15 Experimental determination of Tc from NSCLC adenocarcinoma A549 cell line for *EGFR* exon 21 an exon 19 and *KRAS* exon 2.

Amplicon	T_m (^{o}C)	T_c (^{o}C)	Temperature used for Step 6 in PCR thermocycling (^{o}C)
***EGFR* Exon 19**	86.4	85.6	85.6
***EGFR* Exon 21**	88.4	87.4	87.4
***KRAS* Exon 2**	82.6	81.6	81.0

3.3.2 Confirmation by Agarose Gel Electrophoresis

After the experimental determination of T_c, PCR thermocycling was carried out as described in 2.7.2 with the denaturation temperature in Step 6 set to the appropriate value for each of the exons as per Table 15. Following

thermocycling outlined in Table 6, the PCR products were electrophoresed on a 2% (w/v) agarose gel to confirm amplification (Figures 13 - 15).

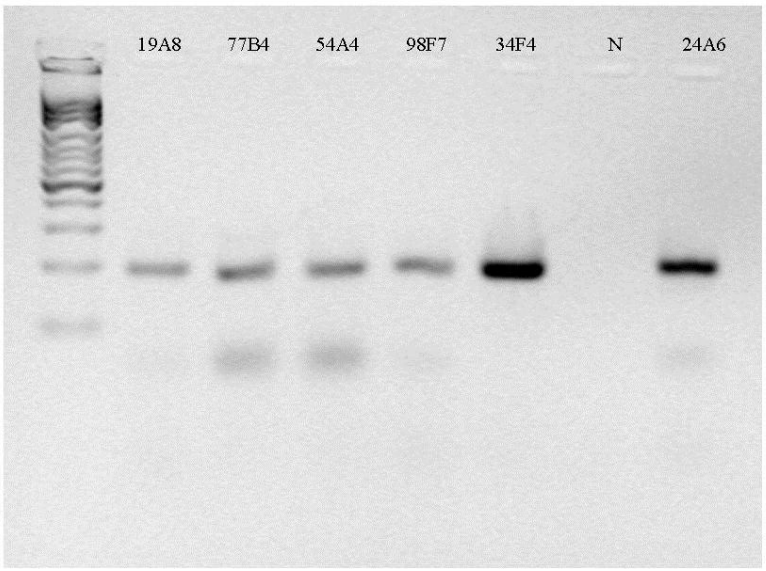

Figure 13 COLD-PCR amplification of *KRAS* exon 2 from genomic DNA isolated from FFPE specimens (Lanes 2-6 and 7), N = No DNA, negative control. Lane 1 = 100bp ladder.

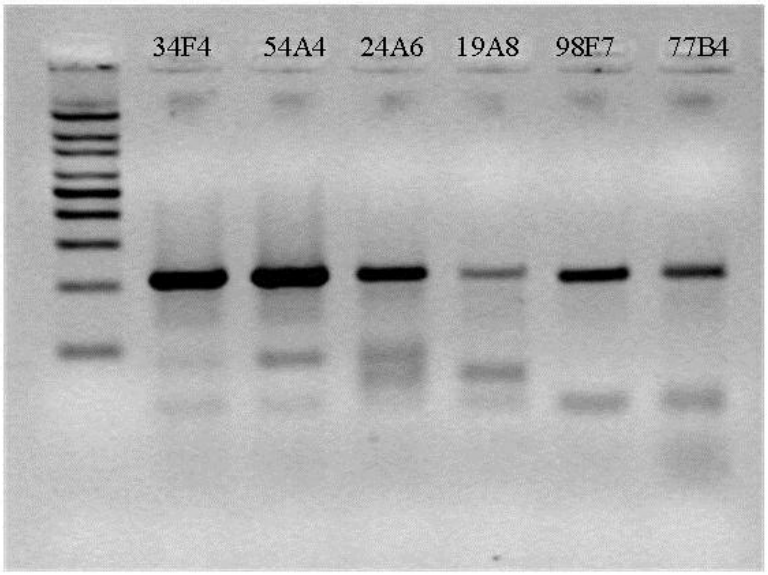

Figure 14 COLD-PCR amplification of *EGFR* exon 21 from genomic DNA isolated from FFPE specimens. Lane 1 = 100bp ladder.

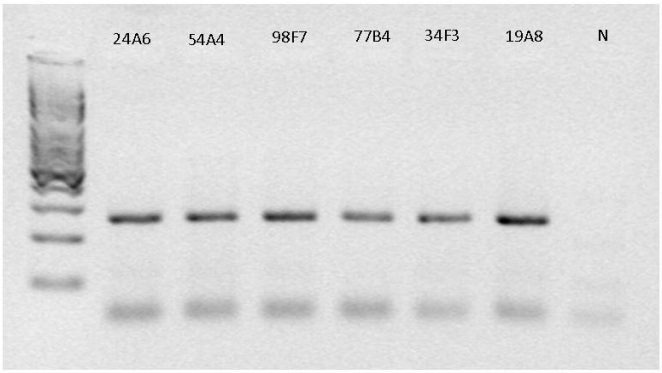

Figure 15 COLD-PCR amplification of EGFR exon 19 from genomic DNA isolated from FFPE specimens (Lanes 2-7), N = no DNA, negative control. Lane 1 = 100bp ladder

3.3.3 Conformation by DNA Sequencing

After migration the COLD-PCR products were purified and sent to the Waikato DNA Sequencing facility for direct sequencing to determine if COLD-PCR could increase the sensitivity of direct sequencing, to enable the detection *EGFR* or *KRAS* mutations in the NSCLC FFPE specimens.

The sequences obtained were entered into 'Genebank' and analysed in BLAST against the reference sequences for *EGFR* exon 21 and 19 and *KRAS* exon 2. Of the sequences obtained from the Waikato Sequencing Facility, none showed any misalignment that would be indicative of the LREA deletion, L858R or a codon 12 or 13 SNP. The sequence data for each of the respective exons from the FFPE specimens is presented in Table 24 in Appendix III. Indeed, this data was interpreted to determine the genotypes for the FFPE specimens, as according to the COLD-PCR and direct sequencing protocol, and presented in Table 16.

Table 16 Genotypes for the FFPE specimens as determined by COLD-PCR and direct sequencing

Specimen	EGFR Exon 21	EGFR Exon 19	KRAS Exon 2
24A6	WT	WT	WT
54A4	WT	WT	WT
98F7	WT	WT	WT
77B4	WT	WT	WT
34F3	WT	WT	WT
19A8	WT	WT	WT

WT = wild-type

3.3.4 Conclusions

Disagreeing with previous reports (Milbury, et al., 2009; Pritchard, et al., 2010; Yu et al., 2011), COLD-PCR was unable to increase the sensitivity of direct sequencing to detect *EGFR* and *KRAS* mutations in the NSCLC FFPE specimens. Indeed, while 'clean' electropherograms were obtained for each of the three exons for all of the FFPE specimens, the inability of this protocol to detect any *EGFR* or *KRAS* mutations was unexpected, given several previous reports of the success of this protocol in increasing the mutated (minor) allele frequency over the 20% that is required for mutation detection using direct sequencing.

Interestingly, specimen 77B4, which was unable to be completely genotyped with our standard PCR and direct sequencing protocol, showed a wild-type sequence for *EGFR* exon 19 in the sequencing data generated using this COLD-PCR protocol. While it was expected that due to the asymmetric amplification of mutated alleles, this specimen would show a LREA mutation in *EGFR* exon 19, given the inherent heterogeneous nature of NSCLC adenocarcinomas, the genomic DNA used this protocol may not have contained sufficient mutated alleles, contributing to this result.

3.4 Restriction-digest mutant enrichment PCR

Restriction-digest mutant enrichment PCR is a way of addressing inability of DNA *Taq* polymerase enzymes to selectively amplify mutant alleles. Indeed, by employing a restriction enzyme to selectively digest the wild-type sequences, the proportion of mutant alleles increases, thereby increasing the amplification of mutant alleles compared with wild-type alleles. This protocol was carried out as per Table 8 using the primers in 10.

3.4.1 Conformation by PAGE Gel Electrophoresis

The PCR products produced from the second round of PCR were electrophoresed on a 13% PAGE gel for restriction fragment length polymorphism analysis. Indeed, Figure 21 shows one of PAGE gel results of this protocol with *EGFR* exon 21. While two product bands can be seen between 300-400bp, this length does not correspond to the bp lengths from the primers as per Table 10. This result depicted in Figure 21 is typical of the results from this protocol with both *EGFR* exon 21 and exon 19, in that we were unable to produce results from which we could accurately genotype the NSCLC FFPE specimens.

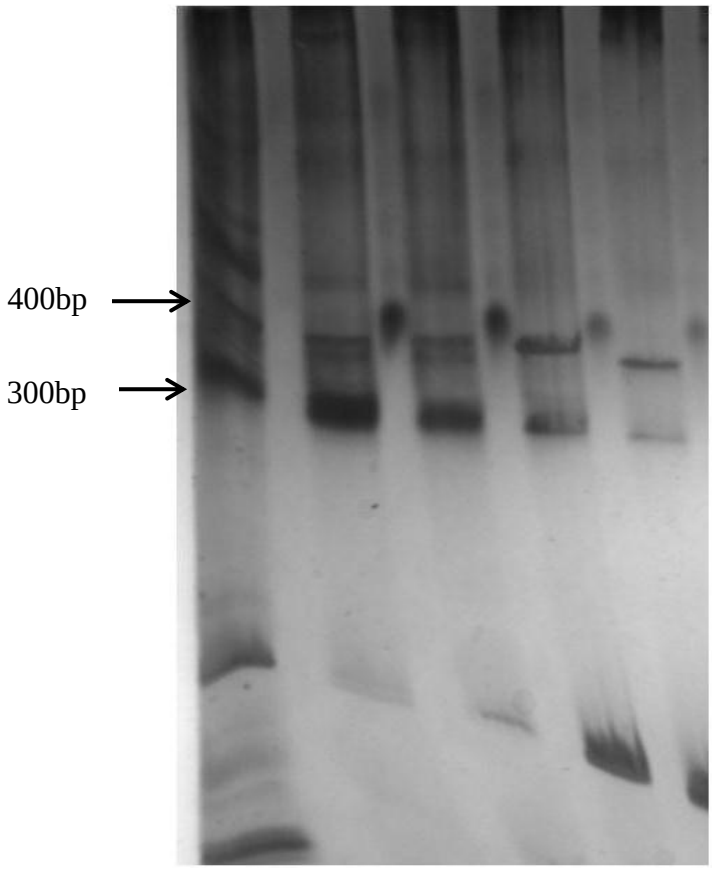

Figure 16 Restriction-digest mutant enrichment PCR 12% PAGE gel from *EGFR* exon 21. Lane 2-5 = FFPE NSCLC specimens. Lane 1 = 100bp ladder. The product sizes seen between 300-400bp did not correspond to the product lengths expected from this protocol.

3.4.2 Conclusions

Given the results in Figure 16, this protocol was not able to genotype the NSCLC specimens, despite the significant modification and manipulation of the variables involved in our protocol outlined in 2.8. Indeed, the digestion incubation lengths, volumes, concentrations as well as the PCR thermocycling conditions were all extensively modified to no avail. As such, we concluded that despite the extensive investigation into the optimum conditions to achieve robust restriction digest mutant enrichment PCR mutation detection assay, we were not able to achieve a robust mutation detection assay using this protocol.

3.5 SSCP-PCR

SSCP-PCR is a gel-based method that has been utilised for the detection of mutations in biopsy specimens for over a decade. Given that silver staining is approximately 100 times more sensitive than ethidium bromide, SSCP-PCR is considered to be a moderately sensitive protocol in that it has shown the ability to detect mutations in a considerable wild-type background. The SSCP protocol was carried out as per Table 5 using the primers in Table 4.

3.5.1 Single Stranded Conformation Polymorphism

The ssDNA products were electrophoresed on a 13% polyacrylamide gel, which was silver stained and the ssDNA banding patterns were compared to those produced the same way from the A549 NSCLC adenocarcinoma cell line. Given that the A549 cell line is known to be wild-type for *EGFR*, any FFPE specimens that showed aberrant banding compared to the A549 ssDNA, where considered indicative of the presence a mutation.

Additionally, genomic DNA was also isolated from the area containing normal tissue from FFPE specimen 34F3 for *KRAS* comparison. This specimen was characterised by an experienced pathologist as containing mostly normal tissue and as a consequence made an ideal source of wild-type *KRAS* ssDNA for identifying *KRAS* mutation in the FFPE specimens.

Indeed, the results of the SSCP-PAGE gel electrophoresis are depicted in Figures 17-21 for the respective exons.

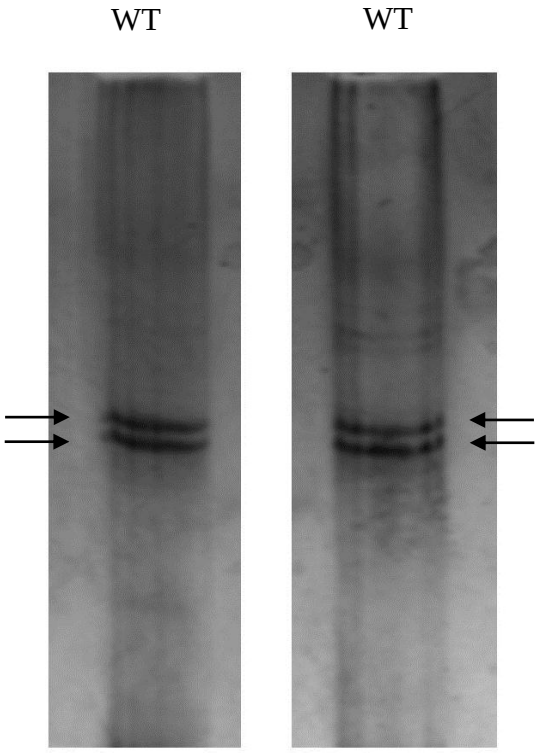

Figure 17 SSCP 13% PAGE gel of *EGFR* exon 19. Left = wild-type A549 NSCLC adenocarcinoma cell line, Right = 34F3 FFPE NSCLC specimen. Banding pattern from the FFPE specimen is the equivalent to that seen from the A549 cell, indicating a wild-type sequence for *EGFR* exon 19 in this specimen.

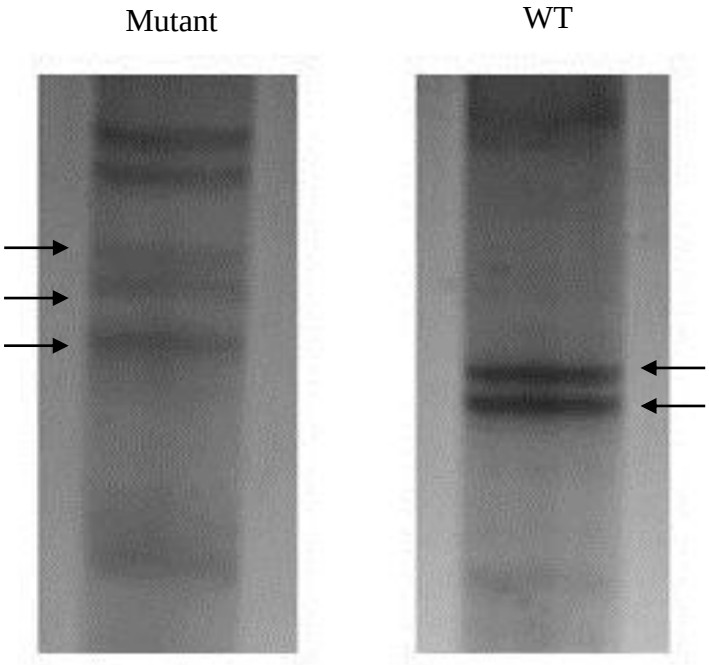

Figure 18 SSCP 13% PAGE gel from *EGFR* exon 19. Left = 54A4 FFPE specimen. Right = wild-type A549 NSCLC adenocarcinoma cell line. Banding pattern from the FFPE specimen shows an aberrant pattern compared to that of the A549 cell, indicating the presence of a mutation in *EGFR* exon 19 in this specimen.

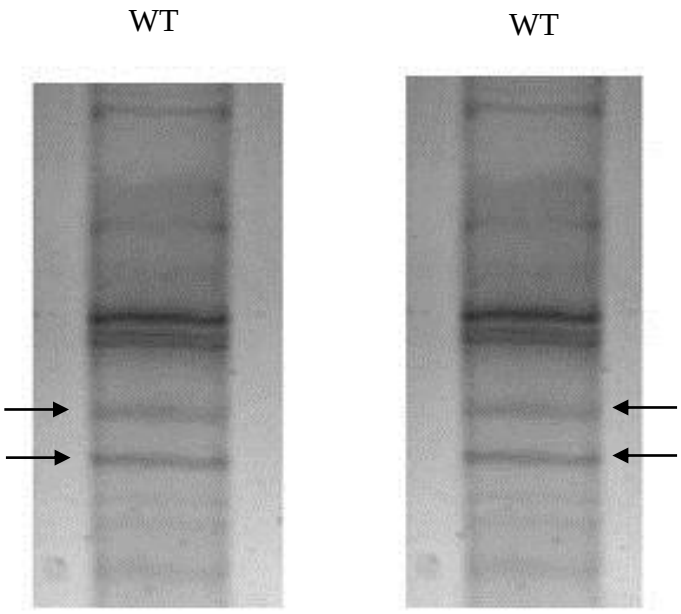

Figure 19 SSCP 13% PAGE gel from *EGFR* exon 21. Left = wild-type A549 NSCLC adenocarcinoma cell line, Right = 54A4 FFPE NSCLC specimen. Banding pattern from the FFPE specimen is equivalent to that seen from the A549 cell, indicating a wild-type sequence for *EGFR* exon 21 in this specimen.

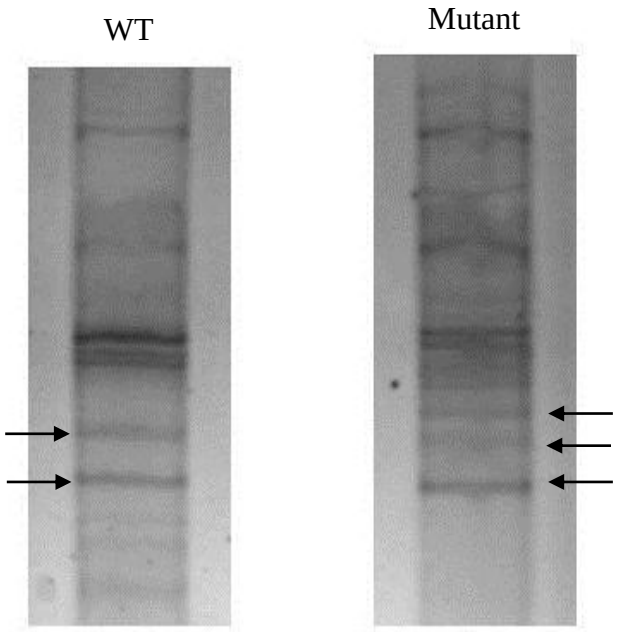

Figure 20 SSCP 13% PAGE gel from EGFR exon 21. Left = A549 NSCLC adenocarcinoma cell line. Right = 98F7 FFPE NSCLC specimen. Banding pattern from the FFPE specimen shows an aberrant pattern compared to that of the A549 cell, indicating the presence of a mutation in *EGFR* exon 21 in this specimen.

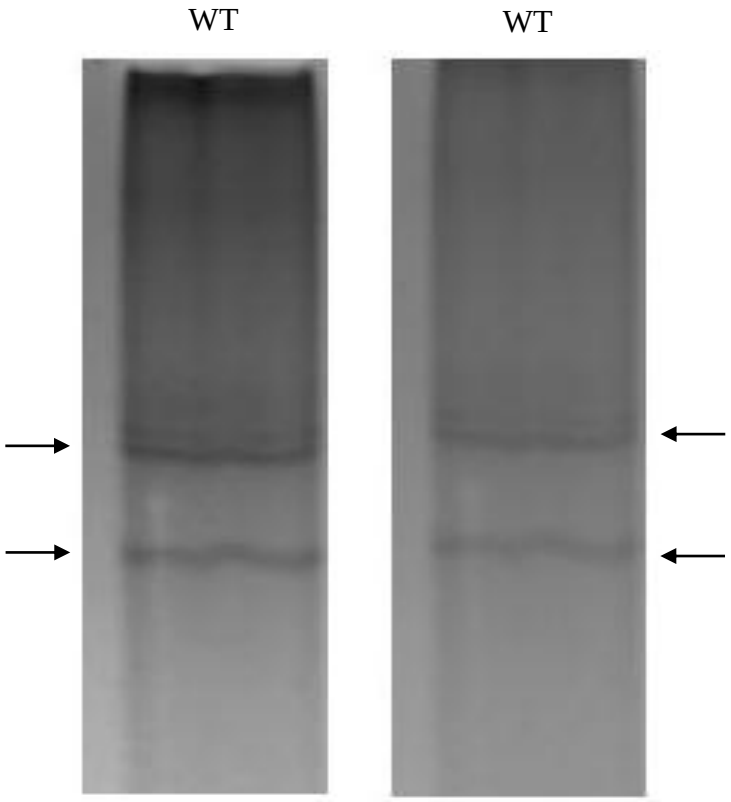

Figure 21 SSCP 13% PAGE gel from *KRAS* exon 2. Left = specimen 77B4 , Right = 34F3 FFPE NSCLC. Banding pattern from the FFPE specimen is equivalent to that seen from the non-tumour genomic DNA isolated from the 34F3 FFPE specimen, indicating a wild-type sequence for the *KRAS* exon 2 in this specimen

Table 17 Genotypes for the seven NSCLC FFPE specimens as determined by SSCP-PCR.

Specimen	EGFR Exon 21	EGFR Exon 19	KRAS Exon 2
24A6	WT	Mutant	WT
54A4	WT	Mutant	WT
98F7	Mutant	WT	WT
77B4	WT	WT	WT
34F3	WT	WT	WT
4674	WT	WT	WT
19A8	-	-	-

WT = wild-type, - = unable to genotype.

3.5.2 Conclusions

SSCP showed promising results as a number of differences in the appearance of bands were identified when comparing the FFPE specimens to those of the wild-type controls. Specifically, specimen 98F7 showed an aberrant pattern compared with the banding pattern produced from the A549 cell line for *EGFR* exon 21, while both specimen 24A6 and 54A4 showed an aberrant banding pattern compared with the banding pattern produced from the A549 cell line for *EGFR* exon 19. No mutations were detected in *KRAS*, however this was not unexpected given that FFPE specimens were selected to enrich for *EGFR* activating mutations and that mutations in *KRAS* are mutually exclusive to those in *EGFR*.

However, while these bands indicate the presence of a mutation in these specimens, direct sequencing is still required to determine the exact position and type of mutation present. Moreover, this protocol was time-consuming in that it required considerable post-PCR processing and relied on a 4hour electrophoresis. Additionally, the inherent inconsistency in silver staining did confound judgement of mutation status of the specimens as there was variability from gel to gel.

3.6 DNA Melt Curves

DNA melt curve analysis has recently been established as a rapid and sensitive method for moderate-throughput and cost-effective screening of mutations in clinical specimens. Indeed, DNA melt curve analysis has been applied to the detection of *EGFR* and *KRAS* mutations in a number of biological specimens. DNA melt curves were produced by heating the products produced from the protocol in Table 12 with the primers in Table 4, at 0.2^{o}C/second.

The resultant melt curves from 2.9.2 were compared to the melt curve produced in the same way from the NSCLC adenocarcinoma A549 cell line. As the mutation status of *EGFR* and *KRAS* is known for these cells, any melt curves produced that showed a left-shifted curve or skewed curves indicated the presence of a mutation. Indeed, the presence of a mutation in the specimen would result in the formation of heteroduplexes, which when compared to homoduplexes of the wild-type, would have lower melting temperature. This results a different melting activity which can be distinguished with the use of an intercalating fluorescent dye.

As such, Figures 22-26 show the melt curves produced for the respective exons using both the A549 cells and the NSCLC FFPE specimens.

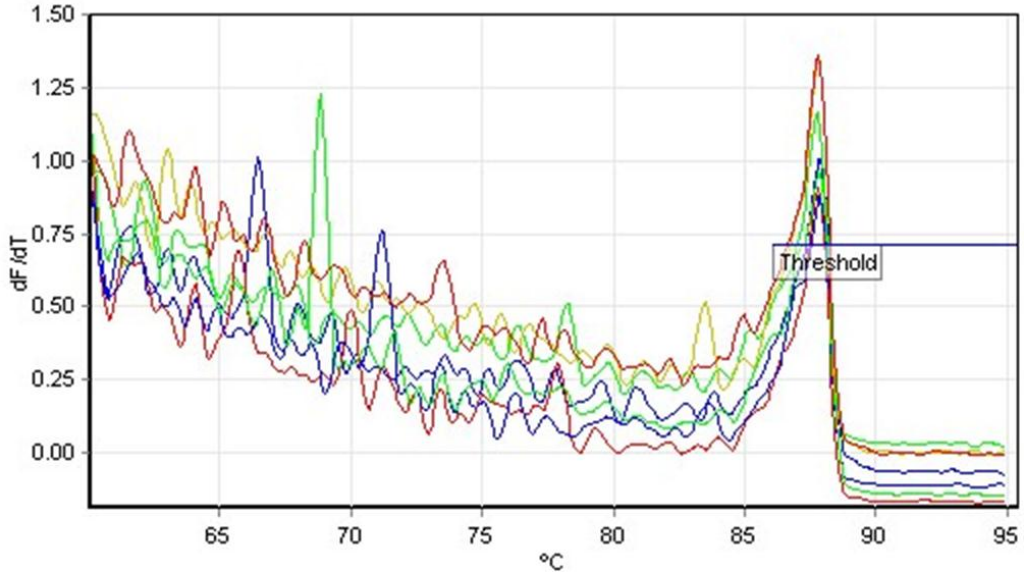

Figure 22 Melt Curve for *EGFR* exon 21 produced from the NSCLC adenocarcinoma A549 cell line. This cell line is known to be wild-type for *EGFR* activating mutations so was used as the control melt curve to which the NSCLC FFPE specimens were compared to.

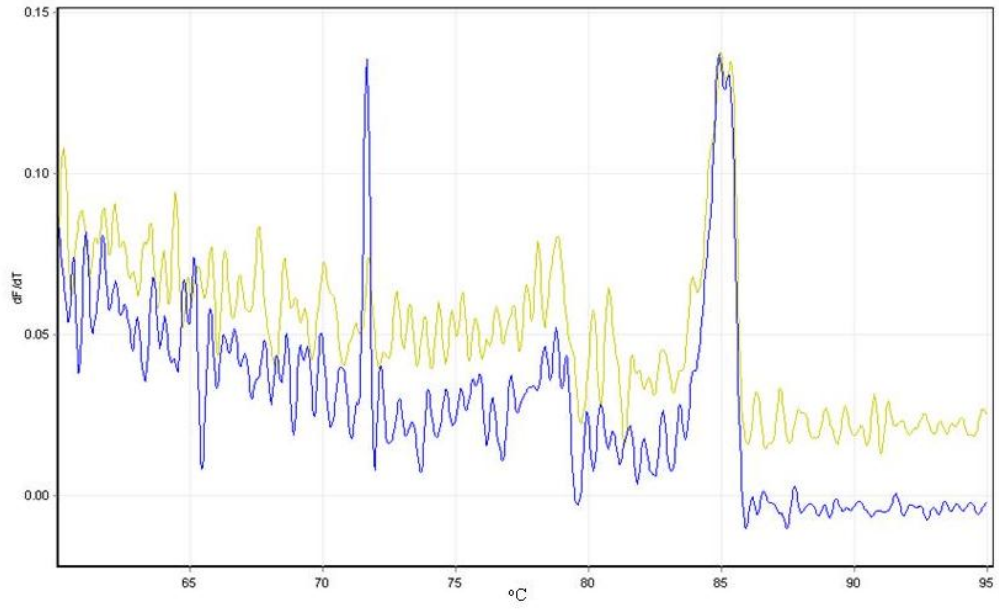

Figure 23 Melt Curve for *EGFR* exon 21 produced from the NSCLE FFPE specimen 98F7. The aberrant melt curve compared to the melt curve produced from the NSCLC adenocarcinoma A549 cell line was indicative of a mutation in *EGFR* exon 21 in this specimen.

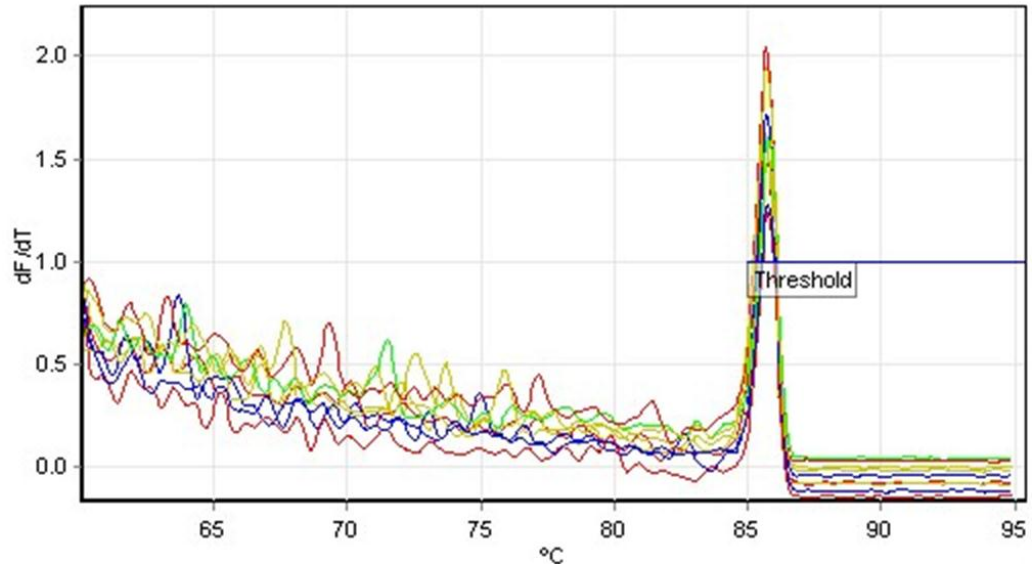

Figure 24 Melt Curve for *EGFR* exon 19 produced from the NSCLC adenocarcinoma A549 cell line. This cell line is known to be wild-type for *EGFR* activating mutations so was used as the control melt curve to which the NSCLC FFPE specimens were compared to.

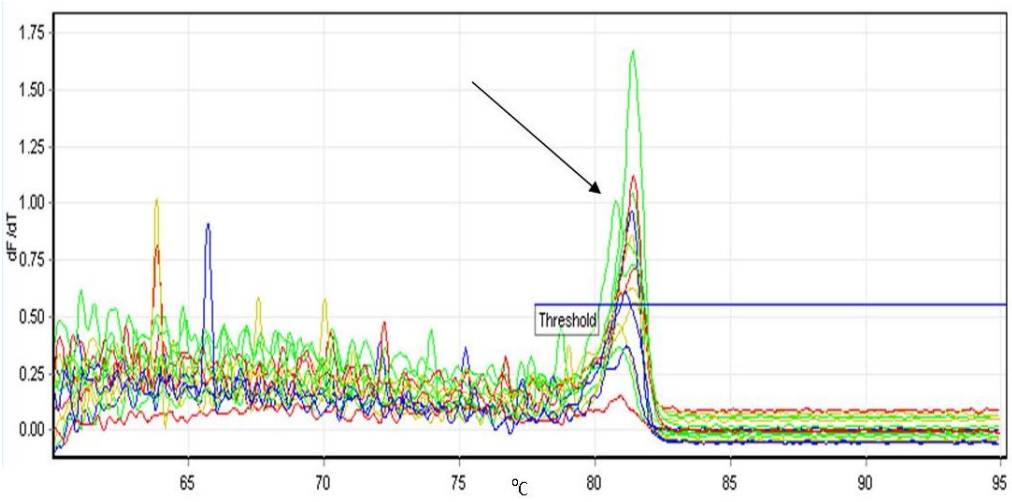

Figure 25 Melt Curve for *EGFR* exon 19 produced from the NSCLC adenocarcinoma A549 cell line and FFPE specimen 54A4 (indicated with arrow). This cell line is known to be wild-type for *EGFR* activating mutations so the aberrant melt curve was indicative of a mutation in this NSCLC FFPE specimens were compared to.

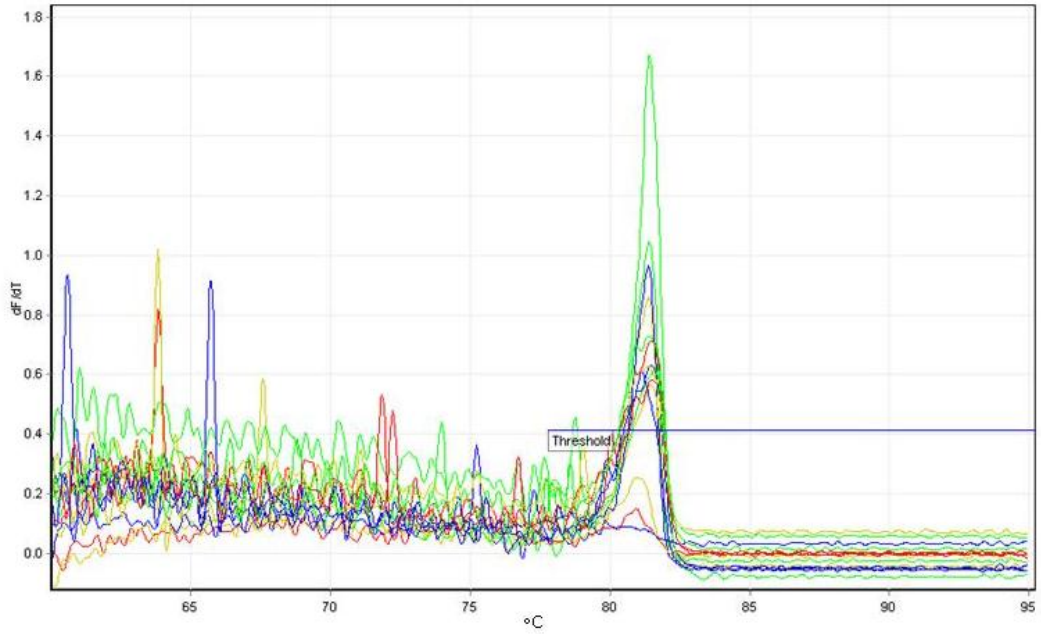

Figure 26 Melt Curve for *KRAS* exon 2 produced from the normal tissue from the NSCLC FFPE specimen 34F3. All other NSCLC FFPE specimens showed the same melt curve profile which was indicative of a wild-type sequence for *KRAS* exon 2.

Table 18 Genotypes for the seven NSCLC FFPE specimens as determined by DNA melt curve analysis

Specimen	EGFR Exon 21	EGFR Exon 19	KRAS Exon 2
24A6	WT	Mutant	WT
54A4	WT	Mutant	WT
98F7	Mutant	WT	WT
77B4	WT	WT	WT
34F3	WT	WT	WT
4674	WT	WT	WT
19A8	-	-	-

WT = wild-type, - = unable to genotype.

3.7 Conclusions

DNA melt curve analysis showed promising results as a number of differences in the melt curves were identified when comparing the FFPE specimens to those of the wild-type controls. Specifically, specimen 98F7 showed an aberrant melt curve compared with the wild-type melt curve produced from the A549 cell line for *EGFR* exon 21, while both specimen 24A6 and 54A4 showed an aberrant melt curve when compared with the melt curve produced from the A549 cell line for EGFR exon 19. No mutations were detected in *KRAS*, however this was not unexpected, as previously mentioned, that FFPE specimens were selected to enrich for *EGFR* activating mutations and that mutations in *KRAS* are mutually exclusive to those in *EGFR*.

While these melt curve indicate the presence of a mutation in these specimens, direct sequencing is still required to determine the exact position and type of mutation present. Moreover, in retrospect, using HRM with a green fluorescent dye such as Syto 13 would have been more applicable, given the greater resolution and therefore sensitivity of this protocol for mutation detection.

3.8 Genotypes by Method

Table 19 Summary of Genotypes for each of the FFPE specimens according to the different mutation detection protocols

Specimen	Standard PCR	COLD-PCR	SSCP	High Resolution DNA Melting Analysis	Physical Characterisation of FFPE specimen
24A6					Mostly tumour tissue
EGFR Exon 21	WT	WT	WT	WT	
EGFR Exon 19	WT	WT	Mutant	Mutant	
KRAS Exon 2	WT	WT	WT	WT	
54A4					Mostly tumour tissue
EGFR Exon 21	WT	WT	WT	WT	
EGFR Exon 19	WT	WT	Mutant	Mutant	
KRAS Exon 2	WT	WT	WT	WT	
98F7					All tumour tissue
EGFR Exon 21	WT	WT	Mutant	Mutant	
EGFR Exon 19	WT	WT	WT	WT	
KRAS Exon 2	WT	WT	WT	WT	
77B4					Mostly tumour tissue
EGFR Exon 21	WT	WT	WT	WT	
EGFR Exon 19	?	WT	WT	WT	
KRAS Exon 2	WT	WT	WT	WT	

34F3					Mostly normal tissue
EGFR Exon 21	WT	WT	WT	WT	
EGFR Exon 19	WT	WT	WT	WT	
KRAS Exon 2	WT	WT	WT	WT	
19A8					Mostly tumour tissue –
EGFR Exon 21	?	WT	-	-	very small specimen
EGFR Exon 19	?	WT	-	-	
KRAS Exon 2	WT	WT	-	-	
4674					
EGFR Exon 21	#	#	WT	WT	Mostly tumour tissue
EGFR Exon 19	#	#	WT	WT	
KRAS Exon 2	#	#	WT	WT	

? = inconclusive results, - = unable to genotype # = protocol was not carried out

3.9 Overall Conclusions

Table 19 represents a summary of the genotypes established using each of the different mutation detection protocols for the seven NSCLC FFPE specimens used in this study. Overall, the sensitivity of these protocols was shown to be rather poor, as only SSCP and DNA melt curve analysis were able to detect putative mutations in the NSCLC FFPE specimens. Specifically, specimens 34F3 and 54A4 were found to harbour an *EGFR* exon 19 mutation while specimen 98F7 was found to harbour an *EGFR* exon 21 mutation. No mutations were detected in *KRAS* exon 2 in the six specimens used for this study. Given that these FFPE specimens were selected due to the possession of predisposing characteristics that would suggest they were likely to harbour mutant positive tumours, our *EGFR* mutation frequency was 43% (n=7), significantly higher than the reported 10% *EGFR* mutation frequency in the Western population.

The two mutant allele enrichment protocols, COLD-PCR and restriction-digest mediated mutant enrichment PCR were implemented with no success. This was unexpected, however false negatives are well-reported given the heterogeneous nature of NSCLC specimens and the degraded DNA template. Also, given that we had such a high yield of genomic DNA from the FFPE specimens, the dilution of this DNA to 100ng/µL for use in PCR may have diluted the mutant concentration to levels below the enrichment capabilities of COLD-PCR and restriction-digest mediated mutant enrichment PCR.

Of the FFPE specimens analysed in this study, 19A8 was particularly problematic as it was a very small specimen which contained a very small proportion of tumour tissue. The genomic DNA isolated from this specimen was minimal and also of relatively poor quality, which limited the inclusion of this specimen in the experimental work. Specifically, I was only able to genotype this specimen using Standard and COLD-PCR followed by direct sequencing due to the lack of DNA extracted from this specimen. As such, while these two protocols showed wild-type sequences for both *EGFR* and *KRAS* genes, I would hesitate to conclude that this specimen does not harbour any mutations given the well-established false negative. Further, of the quality of the DNA isolated from this FFPE specimen was poor, which lead to a number of difficulties and consequently the failure to establish a definite genotype.

Additionally, unlike the other NSCLC FFPE specimens analysed in this study, specimen 4674 was analysed prospectively as this FFPE biopsy specimen was from a current patient and the genotype information elucidated in this study would be used to aid treatment regimen decisions. As such, only SSCP and DNA melt curve analysis were undertaken on this specimen as these mutation detection protocols were shown to have sufficient efficiency to detect *EGFR* mutations in other NSCLC FFPE specimens. Both these protocols found that this specimen was wild-type for both *EGFR* and *KRAS*, which was then reported back to the clinicians to aid their decision making.

Finally, a number of specific conclusions can be made from the experimental work presented in this chapter:

1. Incubation with mineral oil and SDS-lysis solution contributes to a high yield of quality genomic DNA from NSCLC FFPE specimens using a standard Proteinase K and phenol-chloroform extraction.
2. A PCR buffer possessing a high magnesium concentration (4.5mM) is required to produce amplifiable products from FFPE NSCLC specimens in PCR.
3. HOTFIRE *Taq* DNA polymerase is required for efficient amplification from NSCLC FFPE specimens.
4. Standard PCR followed by Direct Sequencing was unable to detect any mutations in the NSCLC FFPE specimens analysed in this study.
5. COLD-PCR followed by direct sequencing was unable to detect any mutations in the NSCLC FFPE specimens analysed in this study.
6. Restriction-digest mediated mutant enrichment PCR involves an extensive and time consuming protocol that requires extensive optimisation and manipulation, which in our hands was unable to produce robust results from which we had the confidence to genotype the NSCLC FFPE specimens.
7. SSCP-PCR is a mutation detection protocol that has sufficient sensitivity to detect *EGFR* mutations in NSCLC FFPE specimens. However, this protocol is labour intensive and has many confounding factors when trying to analyse results.

8. DNA melt curve analysis is a mutation detection protocol that has sufficient sensitivity to detect *EGFR* mutations in NSCLC FFPE specimens. This protocol relatively straightforward to implement as it is an in-tube i.e. non-gel based mutation detection technique.

Chapter 4 General Discussion

Theoretically, the mutant DNA frequency in biological specimens, such as NSCLC FFPE specimens, is usually less than 5% of the wild-type DNA (Akagi et al., 2007). As a consequence, DNA sequencing may not detect the mutant DNA and more sensitive mutation detection techniques must be developed in order to detect mutations in this significant wild-type background. This study illustrates that the detection of *EGFR* and *KRAS* mutations in FFPE NSCLC specimens requires a highly sensitive and robust assay. Indeed, my results, combined with the experimental evidence of others (Beau-Faller et al., 2009; Hlinkova, et al., 2011b; Krypuy, et al., 2006) have shown that direct sequencing, the 'gold standard' mutation detection technique, is sometimes unable to detect the mutation status of *EGFR* and *KRAS* genes in NSCLC FFPE specimens.

Given the clinical significance of these biomarkers for predicting NSCLC patient response to EGFR-TKIs, significant work has been invested in the establishment of a more rapid, sensitive, accurate and cost-effective mutation detection assay. Accordingly, this study also investigated the sensitivity of several alternative PCR-based mutation detection protocols and found of these, SSCP and DNA melt analysis were able to detect *EGFR* mutations in NSCLC FFPE specimens. Although it is difficult to draw strong conclusions from such small sample numbers, this study also found that optimisation of thermocycling conditions is required to produce amplifiable products from FFPE NSCLC specimens.

Further, while technical advancements such as COLD-PCR and restriction digest-mediated mutant enrichment PCR can, through the asymmetric amplification of mutant alleles, theoretically increase the sensitivity of existing DNA analysis techniques, the application of these two protocols was limited in this study. Indeed, the inherent obstacles of FFPE NSCLC specimens, such as the highly heterogeneous nature of NSCLC tumours and the poor quality and quantity of the DNA template in these specimens present significant obstacles that must be overcome.

Moreover, while SSCP and DNA melt analysis can detect putative mutations in FFPE specimens one aspect that remains to be investigated is that conformation of such mutations by direct sequencing. Indeed, both these protocols can detect the presence of a mutation, but they do not provide information on the exact nature and position of the mutation identified. Given the poor sensitivity limit of direct sequencing, such identification is problematic (as shown in this study) as number of mutations can go undetected by direct sequencing. As such, while these methods enable the rapid screening of larger numbers of samples with increased sensitivity, they still require direct sequencing to confirm the clinical significance of the mutation, and further work is required for the establishment of a clinically relevant mutation detection protocol.

In an attempt to develop such a protocol, alternative methods such as amplification refractory mutation systems (ARMS) combined with scorpion method and the peptide nucleic acid-locked (PNA) clamping method, are recently established methods that potentially could be used to rapidly detect rapidly *EGFR* and *KRAS* mutations using real-time PCR–based technology. However, without significant modification these protocols are not applicable in a clinical setting due to the high cost involved. For example the PNAClamp™ EGFR Mutation Detection Kit from Panagene costs US$ 3,750 for 25 tests (personal correspondence, 13/07/2011).

Accordingly, despite significant effort, a routine clinical assay for *EGFR* and *KRAS* mutation detection has yet to be established. While some laboratories offer mutation testing of NSCLC patients, and commercially available kits are becoming readily available, these resources are not often applicable due to the high cost or inaccessibly of the majority of patients to such technology. Given this I investigated the use of direct sequencing, COLD-PCR, restriction-enzyme mediated mutation PCR, single stranded conformation polymorphism and high resolution melting analysis as clinically applicable *EGFR* and *KRAS* mutation detection assays in FFPE NSCLC specimens. This chapter will discuss the key results of this study and evaluate the sensitivity of the PCR-based mutation detection protocols and their suitability for inclusion in clinical practice.

4.1 Isolation of genomic DNA from FFPE Specimens

DNA analysis protocols are dependent on the careful extraction and purification of DNA from FFPE specimens. Consequently, several methods have been established for the extraction of DNA from FFPE specimens and commercial kits are becoming readily available (Miyamae, et al., 2010). However, these extraction methods often require extensive and time consuming protocols which complicate the isolation of quality DNA and limit their inclusion in routine clinical testing (Okello et al., 2010).

DNA extraction and isolation from FFPE specimens is particularly problematic as formalin, whilst an excellent preserver of the integrity of the tissues, compromises the DNA quality, by not only causing the DNA to become fragmented, but also facilitating the formation of cross-links between nucleic acids and proteins. This tends to result in DNA species with average base pair lengths of approximately 200-300bp (Hongdo Do & Dobrovic, 2009). Additionally, the paraffin-embedding process and subsequent storage lead to nucleic acid degradation and modification which further diminish the quality and yield of DNA within these specimens.

To address the problems associated with the extraction of genomic DNA from FFPE specimens, I employed a method adapted from Lin et al. (2009) with considerable success. Specifically, this method utilised mineral oil to deparaffinise the FFPE specimens and incubation with a high pH lysis solution to increase the yield of genomic DNA. After which, the specimen was subjected to a standard extraction method involving Proteinase K and phenol-chloroform. Indeed, the DNA yield from all but one FFPE specimen was significantly higher than anticipated, and in particular, the 260/280 ratios for these specimens were surprisingly close to ideal (as shown in Table 13). Given the low cost and toxicity associated with mineral oil, I concluded that it provides an ideal reagent to extract high yields of quality genomic DNA from FFPE specimens in routine clinical use (J. Lin et al., 2009).

4.2 PCR failure with FFPE DNA template

Most DNA analysis protocols are PCR based and the robust amplification of the genomic DNA isolated from FFPE specimens is crucial for the success of the downstream mutation detection protocol. However, PCR failure is of

particular concern when dealing with FFPE specimens due the chemical processes involved in the fixation and storage process that cause damage to the DNA often resulting in failed reactions (Miyamae, et al., 2010). Indeed, this damage reduces the amplification strength of DNA polymerase enzymes as they often encounter 'blocks' along the FFPE DNA template. Additionally, the miss-incorporation of nucleotides is common with a degraded DNA template, which further reduces the efficiency of amplification, as the enzyme must then replace this via its proofreading and exonuclease activity.

As such, a number of research groups have reported difficulties in working with FFPE specimens, with some even going as far as to suggest that some of the novel variant mutations reported by have been artificially induced by working with small amounts of damaged DNA (Farrugia, et al., 2010). Given that PCR amplification strength from FFPE specimens is adversely related to product size, a typical molecular behaviour of degraded DNA, there is considerable risk associated with the use of DNA isolated from FFPE specimens for mutation detection due to the significant risk of indeterminate or false-negative results (Takano, et al., 2007).

4.2.1 Optimisation of PCR thermocycling conditions

Given the poor amplification ability of DNA isolated from FFPE specimens, significant time and labour went into the optimisation of PCR conditions to not only avoid PCR failure, but also to increase the efficiency of amplification from the FFPE NSCLC specimens. However, despite this considerable effort that went into the optimisation of PCR conditions (as discussed further below), I found that this was not able to overcome the poor sensitivity of direct sequencing.

Significantly, I found that amplification from the A549 NSCLC adenocarcinoma cell line did not require the same optimisation processes, as robust amplification was achieved from the genomic DNA isolated from these cells in thermocycling conditions that were not able to produce amplification products from the FFPE NSCLC specimens. This was not unexpected since the DNA should not be degraded in these cells.

4.2.1.1 Primer design

It has been well documented that extraction from FFPE specimens tends to result in DNA fragments with the average base pair length of 200-300bp (Okello, et al., 2010). In attempt to ensure robust amplification from the FFPE specimens, primers which resulted in amplicon lengths of less than 300bp were selected. Indeed, the resultant PCR products produced from these primers was of suitable length to minimise the adverse effects of formaldehyde on the DNA within the FFPE specimens, and I was able to reproducibly produce amplification products for *EGFR* exon 21 and 19 and *KRAS* exon 2 from all seven specimens.

The primers selected had binding sites within the intro-exon boundary, increasing the likelihood of mutation detection within the exon, as direct sequencing often fails to correctly identify the nucleotides at the ends of the PCR product sequences. Indeed, upon close analysis of the electropherograms, significant background signal could be seen in the beginning and end of all the sequences obtained. Due to this poor signal-to-noise ratio, this part of the sequencing data was removed before BLAST analysis to ensure alignments with the mutation hotspot in the exon.

Moreover, primer selection was also important to ensure the robust nature of the downstream DNA analysis, particularly with SSCP and DNA melt analysis. These two protocols are susceptible to single nucleotide polymorphism (SNP) interference, as they are unable to distinguish between the presence of SNP and a mutation of interest. As such, primers that produced amplification products of approximately 200bp and flanked the exons as closely as possible were selected to reduce the chance of SNP interference confounding results from these protocols.

4.2.1.2 Magnesium Concentration

Magnesium (Mg^{2+}) is an essential ingredient that stabilises the interactions between the oligonucleotide primers, template DNA and *Taq* DNA polymerase

enzyme. Given that Mg^{2+} is a co-factor for the *Taq* DNA polymerase enzyme the Mg^{2+} concentration can greatly affect the efficiency of amplification in PCR. Specifically, it has been shown that an excessive magnesium concentration can cause mis-priming, in that primers bind to incorrect template sites resulting in non-specific products. Insufficient magnesium leads to an inefficient DNA product yield, or in some cases, no production of products at all (Evans, 2009).

While 1.5mM is the standard Mg^{2+} concentration used for most PCR thermocycling, a higher concentration can increase PCR specificity and efficiency (Evans, 2009). Given the poor quality of the template DNA in these experiments, increasing the magnesium was hypothesised to increase the efficiency of amplification, due to the use of a high fidelity *Taq* enzyme with proofreading and exonuclease activity. Indeed, my experiments showed the manipulation of the Mg^{2+} concentration can greatly affect the efficiency of PCR amplification from DNA isolated from FFPE specimens. Specifically, the use of a higher concentration of magnesium (4.5mM) was much more efficient than the standard magnesium concentration of 1.5mM in the production of PCR products from our FFPE specimens. As such, while the standard magnesium concentration was sufficient to produce products from the A549 NSCLC cell line, it was unable to reproducibility produce products from the FFPE specimens.

4.2.1.3 High fidelity DNA *Taq* Polymerase Enzyme

The adverse effects of formalin on the quality of DNA in FFPE specimens are well documented and known to be the source of PCR artefacts. Indeed, when *Taq* DNA polymerase encounters damaged templates it has the propensity to stop nucleotide incorporation and to insert a non-complementary residue (usually an adenosine) into the stand being synthesised (Evans, 2009). As such, a variety of polymerase-mediated sequence changes can occur during PCR amplification, including single base substitutions, deletions and insertions. Of these, single base substitutions that result from the misincorporation of incorrect dNTPs are the most common type of artefactual change (Hongdo Do

& Dobrovic, 2009). This often leads to false positives, as the downstream DNA analysis will detect a change in the DNA sequence and not attribute this to an experimental artefact.

Several researchers have suggested that the detection of mutations, in particular novel variants, in genomic DNA isolated from FFPE specimens is the result of dealing with small quantities of suboptimal quality DNA (J. Lin, et al., 2009). Indeed, more frequent DNA polymerase errors have been detected with direct sequencing when amplifying from FFPE specimens than the matched fresh frozen tissues, highlighting the artificial sequence changes that occur as a result of DNA polymerase error and/or damaged DNA templates (Siwoski, et al., 2002). Do et al. (2009) compared HRM and direct sequencing in FFPE specimens and peripheral blood and while they reported several PCR artefacts with the FFPE specimens, they encountered none with the DNA isolated from the peripheral blood. Additionally, they also investigated if these PCR artefacts they saw in FFPE specimens were caused by DNA polymerase error or chemical damage to the genomic DNA by using two DNA polymerases having different degrees of fidelity. Consequently, they found that PCR artefacts were not reduced as a result of increasing the fidelity of DNA polymerase, leading them to conclude that the damaged DNA template is likely to be the cause of PCR artefacts

As part of this study I investigated the use of two DNA polymerase enzymes and found that HOTFIRE DNA *Taq* Polymerase was required for the robust amplification from genomic DNA isolated from our FFPE specimens. Given that DNA polymerase is a magnesium dependent enzyme, the synergy between the high magnesium buffer concentration and HOTFIRE *Taq* DNA polymerase could have contributed the increased efficiency of amplification from the FFPE specimens and thus this result.

4.3 Direct Sequencing

As previously mentioned, direct sequencing is considered the ‘gold standard’ mutation detection technique however it requires a sufficient amount of tumour tissue of relatively good quality, which is difficult to obtain from NSCLC patients

with inoperable tumours (Do et al 2008). Specifically, it has been shown that the accuracy of this protocol is limited if the biopsy material available contains less than 60% tumour cells (van Zandwijk, et al., 2007). Given that NSCLC specimens often only contain between 2-50% tumour tissue, it has been suggested that most mutations in NSCLC FFPE specimens cannot be detected by direct sequencing (Miyamae, et al., 2010).

Indeed, the low amount of tumour tissue, combined with the inherent heterogeneous nature of NSCLC tumours means that clinically significant somatic mutations such as those in *EGFR* and *KRAS* genes are present at frequencies lower than 25% (given that humans are diploid and oncogenic mutations are often heterozygous). Accordingly, experimental evidence is accumulating that the use of direct sequencing for mutation detection in FFPE specimens results in a number of false negatives, particularly when compared with more sensitive mutation detection techniques (Asano, et al., 2006; Heideman et al., 2009) (discussed further in subsequent sections).

Moreover, due the artificial sequences changes as a result of DNA polymerase error and the degraded FFPE template, the false positive rate of mutation detection with standard PCR and direct sequencing has been established to be approximately 16% in DNA derived from FFPE specimens (Fadhil, Ibrahem, Seth, & Ilyas, 2010). Consequently, there is significant experimental evidence, including the results presented in this study, that highlight the need for the development of alternative protocols which are able to genotype NSCLC FFPE specimens which typically harbour low frequencies of mutated alleles with higher accuracy, selectivity and sensitivity. Additionally, given the high cost, requirement of specialised equipment and the time consuming nature of direct sequencing, there is significant need to develop protocols which are not only more sensitive, but also cost effective, faster and easier to perform, for routine testing to become part of clinical practice.

4.4 COLD-PCR

COLD-PCR is a new form of PCR that has been developed to enrich the mutation-containing alleles through changing the denaturation temperature during

PCR. Indeed, through changing a single parameter of thermocycling, COLD-PCR provides an ideal and simple protocol to increase the sensitivity of PCR-based mutation detection assays (Milbury, et al., 2009). While the COLD-PCR protocol has been used for mutation detection in colorectal and pancreatic cancers (Pritchard, et al., 2010; Yu, et al., 2011), it has not, to my knowledge, been applied to mutation detection in NSCLC.

The application of COLD-PCR to the amplification of low-level *EGFR* and *KRAS* somatic mutations in NSCLC specimens could provide an accurate, convenient and cost-effective way to address the sensitivity and selectivity of PCR-based DNA analysis assays. There are two forms of COLD-PCR, *fast* COLD-PCR and *full* COLD-PCR which differ in terms of their duration and mutation enrichment capabilities. *Full* COLD-PCR takes much longer (approximately 8hours) but is able to detect mutations irrelevant of whether they are known or unknown or what type of mutation they are (transversions or transitions). While, *fast* COLD-PCR takes considerably less time, it is only able to detect known mutations and struggles to detect transitions as these only result in a small change in melting temperature (Pritchard, et al., 2010).

The decreased denaturation temperature of COLD-PCR causes the preferential denaturation of heteroduplexes (formed by hybridisation of mutant and wild-type sequences), and thus enriches mutated allele frequency. As such, COLD-PCR has been reported to increase mutant allelic concentrations sufficient for accurate genotyping via direct sequencing and other mutation detection protocols (Milbury, et al., 2009; Pritchard, et al., 2010; Yu, et al., 2011).

Yu et al. (2011) investigated the use of COLD-PCR to detect *KRAS* mutations in 29 FFPE pancreatic specimens and compared the results with those of standard PCR. Significantly, they found that direct sequencing following standard PCR was only able to detect 11 (37.9%) *KRAS* mutations, while COLD-PCR-mediated direct sequencing detected 21 mutations (72.44%). They also used dilution experiments, whereby they mixed a known concentration of *KRAS* mutant alleles in a background of wild-type alleles, to show that mutation detection with direct sequencing following COLD-PCR has an approximate 5-fold improvement compared with standard PCR.

Pritchard et al. (2010) compared direct sequencing following both standard PCR and COLD-PCR, and HRM analysis following standard PCR and COLD-PCR for the detection of *KRAS* mutations in 61 FFPE colorectal cancer specimens. Of these 61 specimens, they found that all protocols detected *KRAS* mutations in 29 specimens; however, COLD-PCR increased the sensitivity of both HRM analysis and direct sequencing, in that COLD-PCR enhanced HRM and sequencing detected an additional 4 specimens harbouring *KRAS* mutant tumours.

As such, this experimental evidence would suggest that COLD-PCR increases the sensitivity of PCR-based assays to enable accurate detection of low-level mutations such as those present in *EGFR* and *KRAS* in FFPE NSCLC specimens. However, in my hands this protocol, while relatively straightforward, was unable to overcome the sensitivity limit of direct sequencing. Indeed, I was unable to detect any *EGFR* or *KRAS* mutations in our FFPE specimens at all using this protocol, while I was able to detect putative mutations in both SSCP and DNA melt analysis.

4.5 Restriction-enzyme mediated mutant enrichment PCR

Like COLD-PCR, the restriction-enzyme mediated mutant enrichment PCR assay is another protocol that has been established to enrich mutation-containing alleles to levels to enable their detection in significant wild-type background. Specifically, through selection of restriction enzymes with recognition sites exclusively present in the wild-type sequence, the mutant-enriched PCR assay is able to increase the proportion of mutant alleles present in a heterogeneous sample. Thus, during thermocycling mutant alleles will be selectively amplified, contributing to an increase in the sensitivity of downstream mutation detection. Indeed, the use of restriction enzymes to increase low level *EGFR* and *KRAS* mutations has been applied in a number of biological specimens with success.

Specifically, Asano et al (2006) analysed the mutation status of *EGFR* in surgically resected specimens from 108 NSCLC patients using a mutant-enriched PCR assay and compared the results with direct sequencing and non-enriched PCR assay. They found that the mutant-enriched assay was able to detect EGFR mutations in 37 patients, whereas direct sequencing was only able to detect 16

mutations. Given the increased sensitivity of this assay, they concluded that they had developed a highly sensitive mutation detection assay that could be used on clinical samples.

Additionally, Hlinkova et al (2011) analysed 53 archival cytologic specimens that were fixed with methanol and stained with Giemsa or Papanicolaou staining using the same mutant-enrichment protocol developed by Asano et al, to enhance direct sequencing and HRM. They found that direct sequencing was only able to detect five of the 13 *EGFR* mutations identified by HRM. However, when the used mutant-enrichment assay prior to sequencing, all 13 *EGFR* mutations could be detected using this protocol.

Given the previous experimental results highlighting the sensitivity of this protocol to detect *EGFR* mutation in NSCLC patients in a variety of biological tissues, it was hypothesised that it would offer a relatively sensitive and robust mutation detection assay for the use on the FFPE NSCLC specimens used in this study. However, despite significant experimental work, I had very limited success with this protocol, as no robust results could be obtained from our FFPE specimens.

Indeed, significant modification and time-consuming experimental work was carried out to ascertain the optimum conditions to achieve a robust assay, yet I was unsuccessful in elucidating these parameters. The restriction digest incubation often resulted in the 'shearing' of the DNA, that when visualised on an agarose gel, appeared as a giant smear from the well. Conditions for the restriction digest were extensively manipulated, including time, volume, units of enzyme, restriction enzyme buffer volume, BSA concentration, DNA template concentration, yet none of the changes we made to these parameters could establish results with which we could genotype the FFPE specimens.

Given that we used a very similar protocol to both Hlinkova et al and Asano et al, the failure of this protocol was unexpected. However, Asano et al only used computed tomography-guided needle lung biopsies, pleural fluid and surgically resected specimens and Hliknova et al used archival cytologic specimens that were fixed with methanol and stained with Giemsa or Papanicolaou staining. Given the previously described complications as a result of the chemical

processes involved in FFPE biological specimens, I would question if the nature of the genomic DNA isolated from the FFPE NSCLC specimens used in this study was not applicable to this protocol.

As such, I was not able to establish a robust mutant-enrichment protocol, as I had a significant amount of no-results during the experimental work, where the restriction enzymes caused the 'shredding' of the DNA, and smears (no specific bands) were seen when the PCR products were electrophoresed. Additionally, the few times that I was able to produce amplification product there was a significant amount of post-PCR processing involved. Given that this protocol does rely on two rounds of PCR, a restriction digest and the subsequent electrophoresis on a polyacrylamide gel, the rather complex nature of this protocol is not applicable to a high throughput that would be require for it to be applied in clinical practice.

4.6 Single-stranded Conformation Polymorphism

SSCP relies in the principle that the electrophoretic mobility of a ssDNA molecule is highly dependent on its size and structure (Hu, et al., 2007). Given that the ssDNA molecules take up conformations based on their underlying DNA sequence, a mutation at a particular nucleotide position will alter this conformation and hence its migration in a gel. As a consequence, when electrophoresed on a gel matrix, molecules that differ, even as little as a single nucleotide can be distinguished.

Due to the relative simplicity of this protocol, SSCP has been applied to mutation detection in many contexts, including the detection of *EGFR* and *KRAS* mutations in biological specimens. Indeed, Sarkak et al (1995) investigated the use of SSCP to detect *KRAS* mutations in frozen and FFPE tumour specimens of 55 NSCLC patients. They detected 10 of the 55 patients (18%) to have *KRAS* mutations in both frozen and FFPE specimens.

Additionally, Marchetti et al (2005) analysed the tumour and matched normal tissue of 375 NSCLC patients using both SSCP-PCR and direct sequencing. Specimens were taken macroscopically, snap-frozen in liquid nitrogen and stored at -80^{o}C. Using both protocols they found 31 E*GFR* mutations (8%) in the tumour tissue, while no mutations were found in the normal tissue. Indeed, while direct

sequencing could detect *EGFR* mutations in 31 patients, SSCP analysis confirmed these and found an additional eight mutations: seven in exon 21 and one in exon 18.

As such, experimental evidence has led to the conclusion that SSCP-PCR is more sensitive than direct sequencing and can be performed on biological samples, including FFPE specimens, with considerable success (Marchetti et al., 2005; Sarkar et al., 1995). Further, it has also been reported that SSCP requires as little as 10% mutated DNA to detect mutations, which is considerably less than the 20% required by direct sequencing. In accordance with this, I also found that SSCP-PCR was a highly sensitive protocol in that it was able to detect *EGFR* mutations in the NSCLC FFPE specimens used in this study. I found that two specimens 24A6 and 54A4 harboured exon 19 mutations, while specimen 98F7 harboured an exon 21 mutation.

However, while SSCP is more sensitive than direct sequencing, it does require significant post-PCR processing. As such, the production of ssDNA molecules from the PCR products, the lengthy electrophoresis in combination with the labour intensive silver staining procedure, contribute to a significant turn-around time. Given the inherent variability in this process, along with the significant time it takes to obtain results, this protocol is not applicable to clinical practice, given the need for both a high-throughput and rapid turnaround time to aid physician's decision making.

Moreover, since the development of next generation technologies, the reported accuracy of SSCP has since been shown to vary between 60-90% and is highly dependent on several factors that need to be optimised for each area of interest (REF). Moreover, the detection sensitivity decreases further when amplicon length is longer than 200bp, which is smaller than the exons in *EGFR* (Weber, Fukino, Villalona-Calero, & Eng, 2005).

As such, while I have found that this protocol is highly sensitive, there are many aspects of this protocol that make it impractical for application to clinical practice. Given this, further work must be done to establish a protocol that is less time consuming, and requires less post-PCR processing to obtain the same sensitivity and selectivity to detect *EGFR* and *KRAS* mutations in FFPE NSCLC specimens.

4.7 DNA melt analysis and High Resolution DNA melting analysis

Since it was introduced in 2002, HRM has had widespread application to genotyping, mutation scanning and sequence matching (Reed, et al., 2007). As such, HRM analysis is expected to be one of the most practical methods for detecting *EGFR* and *KRAS* mutations in clinical practice. HRM is of particular advantage as it is a closed-tube design whereby an intercalating fluorescent dye is incorporated, without affecting efficiency, during PCR thermocycling. Amplification is immediately followed by the monitoring the change of fluorescence as the intercalating fluorescent dye is released, as the DNA duplexes are slowly heated causing the separation of the two strands (Tindall, Petersen, Woodbridge, Schipany, & Hayes, 2009).

It should be noted that the Corbett instrument used in this study was unable to undertake HRM analysis as this requires the instrument to acquire on the green channel which requires a specific green fluorescent dye such as Syto 13. Indeed, while our protocol was able to detect mutations in the NSCLC analysed, the use of HRM is likely to be more sensitive, and such an area of further research.

One of the major advantages of HRM is that unlike most DNA analysis techniques, no post-PCR processing is required (Reed, et al., 2007). Given that all the previously described protocols require the use of gel-based methods, they are both time consuming and difficult to optimise. As such, in-tube i.e. non-gel protocols such as HRMA are becoming increasing important to increase the simplicity and turnaround time, without comprising sensitivity and accuracy, for the translation of genetic information in clinical practice. Indeed, HRM analysis has been shown to be a highly sensitive non-sequencing method to detect *EGFR* and *KRAS* mutations in a significant wild-type background, such as in FFPE NSCLC specimens.

Takano et al (2007) investigated the use of HRM analysis to detect *EGFR* mutations in 207 NSCLC patients. They also carried out direct sequencing to validate HRM against the ‘gold standard’ mutation detection method. The DNA extracted from FFPE or Papanicolau-stained cytologic specimens was subjected to PCR followed immediately by HRM. The melt curves produced from these

specimens were compared to genomic human DNA, which was used as a control sample to generate melting curves for *EGFR* WT. Those samples that showed skewed or left-shifted curves from the control samples were judged to have a mutation. They found the sensitivity of HRM analysis using DNA extracted from FFPE specimens was 92% and the specificity was 100%, which was significantly higher than direct sequencing. Specifically, *EGFR* exon 19 mutations were detected in 49 (24%) patients and *EGFR* exon 21 mutations were detected in 36 (17%) patients, while the other 122 (59%) were considered to be wild-type.

Additionally, Nomoto et al (2006) used HRMA to detect *EGFR* exon 19 deletions and L858R mutations in archival Papanicolau-stained cytologic specimens from 29 NSCLC patients. They identified *EGFR* mutations in 19 samples and wild-type *EGFR* in 15 samples accurately by HRM analysis, but two samples gave false-negative results and one was indeterminate. As such, the sensitivity of HRM analysis was 88%, while the selectivity was 100%. They also conducted sensitivity studies using 3 adenocarcinoma cell lines, H1650 (delE746-A750), H1975 (L858R mutation) and A549 (WT *EGFR*). Dilutions of *EGFR* mutant cells, H1650 and H1975, where done by using proportions of A549 cells ranging from 100% (no A549 cells) to 0% (no mutant cells). These dilution experiments showed that HRM analysis could detect both *EGFR* exon 21 and exon 19 mutations if at least 10% of cells in a sample where mutants.

Further, Krypuy et al. (2006) investigated the use of HRM analysis to screen 30 NSCLC patients for *KRAS* mutations. Using this protocol they found that 9 of the 30 NSCLC biopsies had *KRAS* mutations and these mutations could also be detected by direct sequencing. They also validated the sensitivity of HRM analysis using dilution experiments, and found that HRM analysis could detect *KRAS* mutations present as low as 6% in wild-type background.

Finally, Farral et al. (2009) combined 19 studies and found the HRM had an overall sensitivity 99.3% (n= 839) and specificity 98.8% (n = 2659). As such, HRM analysis has been shown great efficacy in identifying mutations with less labour, time and expense. Indeed, PCR and melting analysis can be carried out in the same tube within a few hours, and the running cost has been reported to be as low as approximately 1 U.S. dollar per sample (Takano, et al., 2007). HRM analysis enables the high throughput screening of gene mutations, which can

provide clinicians with timely information to aid the selection of appropriate therapeutic choices.

Given the experimental evidence of others, as well as the information presented in this study has highlighted the aptness of analysing the melting behaviour of DNA sequences as a mutation detection technique. Indeed, in my hands DNA melt curve analysis was a highly sensitive protocol in that it was able to detect *EGFR* mutations in the NSCLC FFPE specimens used in this study. Specifically, I found that two specimens 24A6 and 54A4 harboured exon 19 mutations, while specimen 98F7 harboured an exon 21 mutation. As such, I would conclude that the use of DNA melt curve analysis or more specifically, the use of the more sensitive HRM analysis would be an ideal mutation detection protocol to be included in routine clinical practice.

4.8 Clinical implications of developing an assay to detect the mutation status of *EGFR* and *KRAS* genes in NSCLC patients

The mutation status of *EGFR* and *KRAS* is highly predictive of patient response, or lack thereof respectively, to EGFR-TKIs, which have been introduced for the treatment of advanced NSCLC. As such, the development of a clinically relevant assay to detect the mutation status of these two oncogenes in FFPE specimens is crucial for the establishment of personalised medicine in the treatment of NSCLC to ultimately improve patient outcomes and address the poor survival rate associated with the disease.

Currently, aspects such as cost, turnaround time, and quality management are unfavourable with the mutation detection protocols employed. However, the advent of next generation molecular tools provides promising protocols with greater sensitivity, accuracy, precision and selectivity for mutation detection. As such, a clinically relevant mutation detection assay must be easy to apply, rapid, cost-effective, and applicable to small amounts of biopic material and not require extensive labour or sophisticated equipment for the translation of personalised medicine into clinical practice.

Given the inherent obstacles of NSCLC FFPE specimens, further work is required to establish a clinically relevant mutation detection protocol in these patients. As such, while SSCP-PCR and DNA melt curve analysis can detect putative mutations in NSCLC FFPE specimens, the widespread application of these protocols to detect mutations to aid therapeutic treatment decisions is limited. Indeed, I would suggest that until samples from patients are taken specifically for use in genetic testing, the application of personalised medicine in NSCLC is unlikely to occur in clinical practice.

References

Abidoye, O., Ferguson, M. K., & Salgia, R. (2007). Lung carcinoma in African Americans. *Nature clinical practice. Oncology, 4*(2), 118-129.

Akagi, K., Sakai, H., Sudo, J., Hommura, Y., Kurimoto, F., Komagata, H., . . . Yoneda, S. (2007). Simple method to detect important epidermal growth factor receptor gene mutations with bronchoscopic specimens of lung cancer patients for gefitinib treatment. *Targeted Oncology, 2*(3), 145-151.

Allison, M. (2010). Turning the tide in lung cancer. *Nature Biotechnology, 28*(10), 999-1002. doi: 10.1038/nbt1010-999

Ansari, J., Palmer, D. H., Rea, D. W., & Hussain, S. A. (2009). Role of Tyrosine Kinase Inhibitors in Lung Cancer. *Anti-Cancer Agents in Medicinal Chemistry, 9*, 569-575.

Asano, H., Toyooka, S., Tokumo, M., Ichimura, K., Aoe, K., Ito, S., . . . Shimizu, N. (2006). Detection of EGFR gene mutation in lung cancer by mutant-enriched polymerase chain reaction assay. *Clinical Cancer Research, 12*(1), 43-48. doi: 10.1158/1078-0432.ccr-05-0934

Bearz, A., Berretta, M., Lleshi, A., & Tirelli, U. (2011). Target therapies in lung cancer. *Journal of biomedicine & biotechnology, 2011*, 921231-921235.

Beau-Faller, M., Legrain, M., Voegeli, A. C., GuÃ©rin, E., Lavaux, T., Ruppert, A. M., . . . Gaub, M. P. (2009). Detection of K-Ras mutations in tumour samples of patients with non-small cell lung cancer using PNA-mediated PCR clamping. *British journal of cancer, 100*(6), 985-992.

Blakely, T., Shaw, C., Atkinson, J., Cunningham, R., & Sarfati, D. (2011). Social inequalities or inequities in cancer incidence? Repeated census-cancer cohort studies, New Zealand 1981-1986 to 2001-2004. *Cancer Causes & Control, 22*(9), 1307-1318. doi: 10.1007/s10552-011-9804-x

Borras, E., Jurado, I., Hernan, I., Gamundi, M. J., Dias, M., Marti, I., . . . Carballo, M. (2011). Clinical pharmacogenomic testing of KRAS, BRAF and EGFR mutations by high resolution melting analysis and ultra-deep pyrosequencing. [Article]. *Bmc Cancer, 11*, 406-406.

Bunn, J. P. A., Soriano, A., Johnson, G., & Heasley, L. (2000). New therapeutic strategies for lung cancer: biology and molecular biology come of age. *Chest, 117*(4 Suppl 1), 163-168S.

Cappuzzo, F., Hirsch, F. R., Rossi, E., Bartolini, S., Ceresoli, G. L., Bemis, L., . . . Varella-Garcia, M. (2005). Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *Journal of the National Cancer Institute, 97*(9), 643-655. doi: 10.1093/jnci/dji112

Cataldo, V. D., Gibbons, D. L., Perez-Soler, R., & Quintas-Cardama, A. (2011). Treatment of Non-Small-Cell Lung Cancer with Erlotinib or Gefitinib. *New England Journal of Medicine, 364*(10), 947-955. doi: 10.1056/NEJMct0807960

Cho, W. C. S. (2007). Potentially useful biomarkers for the diagnosis, treatment and prognosis of lung cancer. [Article]. *Biomedicine & Pharmacotherapy, 61*(9), 515-519. doi: 10.1016/j.biopha.2007.08.005

Choi, J. E., Park, S. H., Kim, K. M., Lee, W. K., Kam, S., Cha, S. I., . . . Park, J. Y. (2007). Polymorphisms in the epidermal growth factor receptor gene and the risk of primary lung cancer: a case-control study. *Bmc Cancer, 7*. doi: 199
10.1186/1471-2407-7-199

da Cunha Santos, G., Saieg, M. A., Geddie, W., & Leighl, N. (2011). EGFR gene status in cytological samples of nonsmall cell lung carcinoma: controversies and opportunities. *Cancer cytopathology, 119*(2), 80-91.

Dacic, S., Shuai, Y., Yousem, S., Ohori, P., & Nikiforova, M. (2010). Clinicopathological predictors of EGFR/KRAS mutational status in primary lung adenocarcinomas. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc, 23*(2), 159-168.

Do, H., & Dobrovic, A. (2009). Limited copy number-high resolution melting (LCN-HRM) enables the detection and identification by sequencing of low level mutations in cancer biopsies. *Molecular cancer, 8*(1), 82-82.

Do, H., Krypuy, M., Mitchell, P. L., Fox, S. B., & Dobrovic, A. (2008). High resolution melting analysis for rapid and sensitive EGFR and KRAS mutation detection in formalin fixed paraffin embedded biopsies. [Article]. *Bmc Cancer, 8*. doi: 142
10.1186/1471-2407-8-142

Ercan, D., Zejnullahu, K., Yonesaka, K., Xiao, Y., Capelletti, M., Rogers, A., . . . Jänne, P. A. (2010). Amplification of EGFR T790M causes resistance to an irreversible EGFR inhibitor. *Oncogene, 29*(16), 2346-2356.

Evans, M. F. (2009). The polymerase chain reaction and pathology practice. *Diagnostic Histopathology, 15*(7), 344-356.

Fadhil, W., Ibrahem, S., Seth, R., & Ilyas, M. (2010). Quick-multiplex-consensus (QMC)-PCR followed by high-resolution melting: a simple and robust method for mutation detection in formalin-fixed paraffin-embedded tissue. *JOURNAL OF CLINICAL PATHOLOGY, 63*(2), 134-140. doi: 10.1136/jcp.2009.070508

Farrugia, A., Keyser, C., & Ludes, B. (2010). Efficiency evaluation of a DNA extraction and purification protocol on archival formalin-fixed and paraffin-embedded tissue. [Article]. *Forensic Science International, 194*(1-3), E25-E28. doi: 10.1016/j.forsciint.2009.09.004

Fassina, A., Gazziero, A., Zardo, D., Corradin, M., Aldighieri, E., & Rossi, G. P. (2009). Detection of EGFR and KRAS mutations on trans-thoracic needle aspiration of lung nodules by high resolution melting analysis. [Article]. *JOURNAL OF CLINICAL PATHOLOGY, 62*(12), 1096-1102. doi: 10.1136/jcp.2009.067587

Franklin, W. A., Haney, J., Sugita, M., Bemis, L., Jimeno, A., & Messersmith, W. A. (2010). KRAS Mutation. *The Journal of Molecular Diagnostics, 12*(1), 43-50.

Fukui, T., Ohe, Y., Tsuta, K., Furuta, K., Sakamoto, H., Takano, T., . . . Tamura, T. (2008). Prospective study of the accuracy of EGFR mutational analysis by high-resolution melting analysis in small

samples obtained from patients with non-small cell lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research, 14*(15), 4751-4757.

Furuta, K., Yokozawa, K., Matsuno, Y., Nomoto, K., Ohe, Y., Maeshima, A. M., . . . Shibata, T. (2006). Detection of EGFR Mutations in Archived Cytologic Specimens of Non-Small Cell Lung Cancer Using High-Resolution Melting Analysis. *American Journal of Clinical Pathology, 126*(4), 608-615.

Gazdar, A. F. (2010). Epidermal growth factor receptor inhibition in lung cancer: the evolving role of individualized therapy. *Cancer and Metastasis Reviews, 29*(1), 37-48. doi: 10.1007/s10555-010-9201-z

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell, 100*(1), 57-70.

Hann, C. L., & Brahmer, J. R. (2007). Who should receive epidermal growth factor receptor inhibitors for non-small cell lung cancer and when? *Current treatment options in oncology, 8*(1), 28-37. doi: 10.1007/s11864-007-0024-2

He, C., Liu, M., Zhou, C., Zhang, J., Ouyang, M., Zhong, N., & Xu, J. (2009). Detection of' epidermal growth factor receptor mutations in plasma by mutant-enriched PCR assay for prediction of the response to gefitinib in patients with non-small-celi lung cancer. *International Journal of Cancer, 125*(10), 2393-2399. doi: 10.1002/ijc.24653

Heideman, D. A. M., Thunnissen, F. B., Doeleman, M., Kramer, D., Verheul, H. M., Smit, E. F., . . . Snijders, P. J. F. (2009). A panel of high resolution melting (HRM) technology-based assays with direct sequencing possibility for effective mutation screening of EGFR and K-ras genes. [Article]. *Cellular Oncology, 31*(5), 329-333. doi: 10.3233/clo-2009-0489

Herbst, R. S., Heymach, J. V., & Lippman, S. M. (2008). Molecular origins of cancer: Lung cancer. [Review]. *New England Journal of Medicine, 359*(13), 1367-1380. doi: 10.1056/NEJMra0802714

Hirsch, F. R. (2009). The role of genetic testing in the prediction of response to EGFR inhibitors in NSCLC. *Oncogene, 28*(31), S1(3).

Hlinkova, K., Babal, P., Berzinec, P., Majer, I., & Ilencikova, D. (2011a). Rapid and Efficient Detection of EGFR Mutations in Problematic Cytologic Specimens by High-Resolution Melting Analysis. *Molecular Diagnosis & Therapy, 15*(1), 21-29.

Hlinkova, K., Babal, P., Berzinec, P., Majer, I., & Ilencikova, D. (2011b). Rapid and efficient detection of EGFR mutations in problematic cytologic specimens by high-resolution melting analysis. *Molecular Diagnosis & Therapy, 15*(1), 21-29.

Hodgen, E., Tobias, M., & Chenung, J. (2002). Cancer in New Zealand Trends and Projections: A Summary: Ministry of Health.

Hu, M., Chilton, N. B., Campbell, B. E., Jex, A. J., Otranto, D., Cafarchia, C., . . . Zhu, X. (2007). Single-strand conformation polymorphism (SSCP) for the analysis of genetic variation. *Nature Protocols, 1*(6), 3121-3128.

Igbokwe, A., & Lopez-Terrada, D. H. (2011). Molecular testing of solid tumors. *Archives of pathology & laboratory medicine, 135*(1), 67-82.

Jacobson, G. (1998). *Novel clinical DNA sources for p53 gene analysis: An investigation of colorectal cancers and mesotheliomas.* Master of Science, The University of Waikato, Hamilton.

Jančík, S., Drábek, J., Radzioch, D., & Hajdúch, M. (2010). Clinical relevance of KRAS in human cancers. *Journal of biomedicine & biotechnology, 2010*(Journal Article), 150960-151509.

Jemal, A., Siegel, R., Xu, J. Q., & Ward, E. (2010). Cancer Statistics, 2010. [Article]. *Ca-a Cancer Journal for Clinicians, 60*(5), 277-300. doi: 10.1002/caac.20073

Jimeno, A., & Hidalgo, M. (2006). Pharmacogenomics of epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors. *BIOCHIMICA ET BIOPHYSICA ACTA-REVIEWS ON CANCER, 1766*(2), 217-229.

John, T., Liu, G., & Tsao, M. S. (2009). Overview of molecular testing in non-small-cell lung cancer: mutational analysis, gene copy number, protein expression and other biomarkers of EGFR for the prediction of response to tyrosine kinase inhibitors. *Oncogene, 28*(S1), S14-S23.

Johnson, B. E., & Kelley, M. J. (1995). Biology of Small Cell Lung Cancer. [Article; Proceedings Paper]. *Lung Cancer, 12*, S5-S16. doi: 10.1016/s0169-5002(10)80014-5

Kim, E. S., Herbst, R. S., Wistuba, II, Lee, J. J., Blumenschein, G. R., Tsao, A., . . . Hong, W. K. (2011). The BATTLE Trial: Personalizing Therapy for Lung Cancer. [Article]. *Cancer Discovery, 1*(1), 44-53. doi: 10.1158/2159-8274.cd-10-0010

Kim, T.-Y., Han, S.-W., & Bang, Y.-J. (2007). Chasing targets for EGFR tyrosine kinase inhibitors in non-small-cell lung cancer: Asian perspectives. *Expert Rev. Mol. Diagn, 7*(6), 821-836.

Kolch, W., & Pitt, A. (2010). Functional proteomics to dissect tyrosine kinase signalling pathways in cancer. *Nature Reviews Cancer, 10*, 618-629.

Kosaka, T., Yatabe, Y., Endoh, H., Kuwano, H., Takahashi, T., & Mitsudomi, T. (2004). Mutations of the epidermal growth factor receptor gene in lung cancer: Biological and clinical implications. *Cancer Research, 64*(24), 8919-8923. doi: 10.1158/0008-5472.can-04-2818

Kristensen, L. S., Daugaard, I. L., Christensen, M., Hamilton-Dutoit, S., Hager, H., & Hansen, L. L. (2010). Increased Sensitivity of KRAS Mutation Detection by High-Resolution Melting Analysis of COLD-PCR Products. *Human Mutation, 31*(12), 1366-1373. doi: 10.1002/humu.21358

Krypuy, M., Newnham, G. M., Thomas, D. M., Conron, M., & Dobrovic, A. (2006). High resolution melting analysis for the rapid and sensitive detection of mutations in clinical samples: KRAS codon 12 and 13 mutations in non-small cell lung cancer. [Article]. *Bmc Cancer, 6*. doi: 295

10.1186/1471-2407-6-295

Ladanyi, M., & Pao, W. (2008). Lung adenocarcinoma: guiding EGFR-targeted therapy and beyond. [Article]. *Modern Pathology, 21*, S16-S22. doi: 10.1038/modpathol.3801018

Lee, Y. J. L. Y. J., Kim, J. H., Kim, S. K., Ha, S. J., Mok, T. S., Mitsudomi, T., & Cho, B. C. (2011). Lung cancer in never smokers: Change of a mindset in the molecular era. [Review]. *Lung Cancer, 72*(1), 9-15. doi: 10.1016/j.lungcan.2010.12.013

Li, J., & Makrigiorgos, M. (2009). COLD-PCR: a new platform for highly improves mutation detection in cancer and genetic testing. *Biochemical Society Transactions, 37*, 427-432.

Li, J., Wang, L., Mamon, H., Kulke, M. H., Berbeco, R., & Makrigiorgos, G. M. (2008). Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nature Medicine, 14*(5), 579-584. doi: 10.1038/nm1708

Lin, C.-C., & Yang, J. C.-H. (2011). Optimal management of patients with non-small cell lung cancer and epidermal growth factor receptor mutations. *Drugs, 71*(1), 79-88.

Lin, C. H., Yeh, K. T., Chang, Y. S., Hsu, N. C., & Chang, J. G. (2010). Rapid detection of epidermal growth factor receptor mutations with multiplex PCR and primer extension in lung cancer. [Article]. *Journal of Biomedical Science, 17*. doi: 37

10.1186/1423-0127-17-37

Lin, J., Kennedy, S. H., Svarovsky, T., Rogers, J., Kemnitz, J. W., Xu, A., & Zondervan, K. T. (2009). High-quality genomic DNA extraction from formalin-fixed and paraffin-embedded samples deparaffinized using mineral oil. *Analytical Biochemistry, 395*(2), 265-267. doi: 10.1016/j.ab.2009.08.016

Luthra, R., & Zuo, Z. (2009). COLD-PCR finds hot application in mutation analysis. *Clinical chemistry, 55*(12), 2077-2078.

Ma, B. B. Y., Hui, E. P., & Mok, T. S. K. (2010). Population-based differences in treatment outcome following anticancer drug therapies. *Lancet Oncology, 11*(1), 75-84.

Mancl, E. E., Kolesar, J. M., & Vermeulen, L. C. (2009). Clinical and economic value of screening for Kras mutations as predictors of response to epidermal growth factor receptor inhibitors. *American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists, 66*(23), 2105-2112.

Marchetti, A., Martella, C., Felicioni, L., Barassi, F., Salvatore, S., Chella, A., . . . Buttitta, F. (2005). EGFR mutations in non-small-cell lung cancer: Analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment. [Article]. *Journal of Clinical Oncology, 23*(4), 857-865. doi: 10.1200/jco.2005.08.043

Masago, K., Fujita, S., Mio, T., Ichikawa, M., Sakuma, K., Kim, Y. H., . . . Mishima, M. (2008). Accuracy of epidermal growth factor receptor gene mutation analysis by direct sequencing method based on small biopsy specimens from patients with non-small cell lung cancer: analysis of results in 19 patients. *International Journal of Clinical Oncology, 13*(5), 442-446.

Milbury, C. A., Li, J., & Makrigiorgos, G. M. (2009). COLD-PCR-Enhanced High-Resolution Melting Enables Rapid and Selective Identification of Low-Level Unknown Mutations. [Article]. *Clinical chemistry, 55*(12), 2130-2143. doi: 10.1373/clinchem.2009.131029

Mitsudomi, T. (2010). Advances in Target Therapy for Lung Cancer. *Japanese Journal of Clinical Oncology, 40*(2), 101-106. doi: 10.1093/jjco/hyp174

Miyamae, Y., Shimizu, K., Mitani, Y., Araki, T., Kawai, Y., Baba, M., . . . Takeyoshi, I. (2010). Mutation Detection of Epidermal Growth Factor Receptor and KRAS Genes Using the Smart Amplification Process Version 2 from Formalin-Fixed, Paraffin-Embedded Lung Cancer Tissue. [Article]. *Journal of Molecular Diagnostics, 12*(2), 257-264. doi: 10.2353/jmoldx.2010.090105

Mok, T. (2009). Front line studies using clinical parameters for EGFR-TKI. [Meeting Abstract]. *JOURNAL OF THORACIC ONCOLOGY, 4*(9), S76-S77.

Morlan, J., Baker, J., & Sinicropi, D. (2009). Mutation Detection by Real-Time PCR: A Simple, Robust and Highly Selective Method. *Plos One, 4*(2). doi: e4584

10.1371/journal.pone.0004584

Nomoto, K., Tsuta, K., Takano, T., Fukui, T., Yokozawa, K., Sakamoto, H., . . . Matsuno, Y. (2006). Detection of EGFR mutations in archived cytologic specimens of non-small cell lung cancer using high-resolution melting analysis. *American journal of clinical pathology, 126*(4), 608-615.

O'Hagan, R. C., & Heyer, J. (2011). KRAS Mouse Models: Modeling Cancer Harboring KRAS Mutations: Modeling Cancer Harboring KRAS Mutations. *Genes & Cancer, 2*(3), 335-343.

Okello, J. B. A., Zurek, J., Devault, A. M., Kuch, M., Okwi, A. L., Sewankambo, N. K., . . . Poinar, H. N. (2010). Comparison of methods in the recovery of nucleic acids from archival formalin-fixed paraffin-embedded autopsy tissues. *Analytical Biochemistry, 400*(1), 110-117. doi: 10.1016/j.ab.2010.01.014

Pao, W., & Girard, N. (2010). New driver mutations in non-small-cell lung cancer. [Review]. *Lancet Oncology, 12*(2), 175-180. doi: 10.1016/s1470-2045(10)70087-5

Pao, W., Iafrate, A. J., & Su, Z. L. (2011). Genetically informed lung cancer medicine. [Review]. *Journal of Pathology, 223*(2), 230-240. doi: 10.1002/path.2788

Pao, W., Kris, M. G., Iafrate, A. J., Ladanyi, M., Janne, P. A., Wistuba, II, . . . Lynch, T. J. (2009). Integration of Molecular Profiling into the Lung Cancer Clinic. [Article]. *Clinical Cancer Research, 15*(17), 5317-5322. doi: 10.1158/1078-0432.ccr-09-0913

Pao, W., Wang, T. Y., Riely, G. J., Miller, V. A., Pan, Q. L., Ladanyi, M., . . . Varmus, H. E. (2005). KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. [Article]. *Plos Medicine, 2*(1), 57-61. doi: e17

10.1371/journal.pmed.0020017

Penzel, R., Sers, C., Chen, Y., Lehmann-Mühlenhoff, U., Merkelbach-Bruse, S., Jung, A., . . . Schirmacher, P. (2011). EGFR mutation detection in NSCLC-assessment of diagnostic application and recommendations of the German Panel for Mutation Testing in NSCLC. *Virchows Archiv, 458*(1), 95-98.

Pierre-Jean, L., & William, J. (2012). Worldwide variations in EGFR somatic mutations: a challenge for personalized medicine. *Diagnostic Pathology, 7*(1), 13. doi: 10.1186/1746-1596-7-13

Pritchard, C. C., Akagi, L., Reddy, P. L., Joseph, L., & Tait, J. F. (2010). COLD-PCR enhanced melting curve analysis improves diagnostic accuracy for KRAS mutations in colorectal carcinoma. *BMC clinical pathology, 10*, 6. doi: 10.1186/1472-6890-10-6

Provencio, M., Garcia-Campelo, R., Isla, D., & de Castro, J. (2009). Clinical-molecular factors predicting response and survival for tyrosine-kinase inhibitors. [Article]. *Clinical & Translational Oncology, 11*(7), 428-436. doi: 10.1007/s12094-009-0381-3

Ramalingam, S. S., Owonikoko, T. K., & Khuri, F. R. (2011). Lung Cancer: New Biological Insights and Recent Therapeutic Advances. *Ca-a Cancer Journal for Clinicians, 61*(2), 91-112. doi: 10.3322/caac.20102

Reed, G. H., Kent, J. O., & Wittwer, C. T. (2007). High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Pharmacogenomics, 8*(6), 597-608.

Rosell, R., Ichinose, Y., Taron, M., Sarries, C., Queralt, C., Mendez, P., . . . Sanchez, J. J. (2005). Mutations in the tyrosine kinase domain of the EGFR gene associated with gefitinib response in non-small-cell lung cancer. [Article]. *Lung Cancer, 50*(1), 25-33. doi: 10.1016/j.lungcan.2005.05.017

Rosell, R., Moran, T., Cardenal, F., Porta, R., Viteri, S., Angel Molina, M., . . . Taron, M. (2010). Predictive biomarkers in the management of EGFR mutant lung cancer. In D. Braaten (Ed.), (Vol. 1210, pp. 45-52).

Rosell, R., Moran, T., Queralt, C., Porta, R., Cardenal, F., Camps, C., . . . Spanish Lung Canc, G. (2009). Screening for Epidermal Growth Factor Receptor Mutations in Lung Cancer. [Article]. *New England Journal of Medicine, 361*(10), 958-U938. doi: 10.1056/NEJMoa0904554

Santarpia, M., Salazar, M. F., Pettineo, G., Magri, I., Benecchi, S., Altavilla, G., & Rosell, R. (2011). Tyrosine kinase inhibitors for non-small-cell lung cancer: finding patients who will be responsive. *Expert Review of Respiratory Medicine, 5*(3), 413-424.

Sarkar, F. H., Valdivieso, M., Borders, J., Yao, K. L., Raval, M. M. T., Madan, S. K., . . . Crissman, J. D. (1995). A UNIVERSAL METHOD FOR THE MUTATIONAL ANALYSIS OF K-RAS AND P53 GENE IN NON-SMALL-CELL LUNG-CANCER USING FORMALIN-FIXED PARAFFIN-EMBEDDED TISSUE. [Article]. *Diagnostic Molecular Pathology, 4*(4), 266-273. doi: 10.1097/00019606-199512000-00007

Schubbert, S., Shannon, K., & Bollag, G. (2007). Hyperactive Ras in developmental disorders and cancer. *Nature Reviews Cancer, 7*(4), 295-308. doi: 10.1038/nrc2109

Sequist, L. V., Bell, D. W., Lynch, T. J., & Haber, D. A. (2007). Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *Journal of Clinical Oncology, 25*(5), 587-595. doi: 10.1200/jco.2006.07.3585

Sequist, L. V., & Lynch, T. J. (2008). EGFR tyrosine kinase inhibitors in lung cancer: an evolving story. *Annual review of medicine, 59*(1), 429-442.

Sharma, S. V., Bell, D. W., & Settleman, J. (2007). Epidermal growth factor receptor mutations in lung cancer. *Nature Reviews Cancer, 7*(3), 169-181.

Siwoski, A., Ishkanian, A., Garnis, C., Zhang, L. W., Rosin, M., & Lam, W. L. (2002). An efficient method for the assessment of DNA quality of archival microdissected specimens. [Article]. *Modern Pathology, 15*(8), 889-892. doi: 10.1097/01.mp.0000024288.63070.4f

Squassina, A., Manchia, M., Manolopoulos, V. G., Artac, M., Lappa-Manakou, C., Karkabouna, S., . . . Pantrinos, G. P. (2010). Realities and expectations of pharmacogenomics and personalized medicine: impact of translating genetic knowledge into clinical practice. [Article]. *Pharmacogenomics 11*(8), 1149-1167. doi: 10.2217/pgs.10.97

Sriram, K. B., Larsen, J. E., Yang, I. A., Bowman, R. V., & Fong, K. M. (2011). Genomic medicine in non-small cell lung cancer: Paving the path to personalized care. *Respirology, 16*(2), 257-263. doi: 10.1111/j.1440-1843.2010.01892.x

Stevens, W., Stevens, G., Kolbe, J., & Cox, B. (2008). Management of stages I and II non-small-cell lung cancer in a New Zealand study: divergence from international practice and recommendations. *Internal medicine journal, 38*(10), 758-768.

Suda, K., Tomizawa, K., & Mitsudomi, T. (2010). Biological and clinical significance of KRAS mutations in lung cancer: an oncogenic driver that contrasts with EGFR mutation. *Cancer metastasis reviews, 29*(1), 49-60.

Sun, S., Schiller, J. H., & Gazdar, A. F. (2007). Lung cancer in never smokers--a different disease. *Nature reviews. Cancer, 7*(10), 778-790.

Sutherland, T. J. T., & Aitken, D. (2008). Ethnic and socioeconomic inequalities in lung cancer in a New Zealand population. *Respirology, 13*(4), 590-593.

Takano, T., Ohe, Y., Tsuta, K., Fukui, T., Sakamoto, H., Yoshida, T., . . . Tamura, T. (2007). Epidermal growth factor receptor mutation detection using high-resolution melting analysis predicts outcomes in patients with advanced non-small cell lung cancer treated with gefitinib. *Clinical Cancer Research, 13*(18), 5385-5390. doi: 10.1158/1078-0432.ccr-07-0627

Takeuchi, K., & Ito, F. (2010). EGF receptor in relation to tumor development: molecular basis of responsiveness of cancer cells to EGFR-targeting tyrosine kinase inhibitors. *Febs Journal, 277*(2), 316-326. doi: 10.1111/j.1742-4658.2009.07450.x

Tindall, E. A., Petersen, D. C., Woodbridge, P., Schipany, K., & Hayes, V. M. (2009). Assessing High-Resolution Melt Curve Analysis for Accurate Detection of Gene Variants in Complex DNA Fragments. [Article]. *Human Mutation, 30*(6), 876-883. doi: 10.1002/humu.20919

Triano, L. R., Deshpande, H., & Gettinger, S. N. (2010). Management of patients with advanced non-small cell lung cancer: current and emerging options: Current and Emerging Options. *Drugs, 70*(2), 167.

Uhara, M., Matsuda, K., Taira, C., Higuchi, Y., Okumura, N., & Yamauchi, K. (2009). Simple polymerase chain reaction for the dectection of mutations and deletions in the epidermal growth factor receptor gene: Applications of this method for the diagnosis of non-small-cell lung cancer. *Clinica Chimica Acta, 401*, 68-72.

van Zandwijk, N., Mathy, A., Boerrigter, L., Ruijter, H., Tielen, I., de Jong, D., . . . Nederlof, P. (2007). EGFR and KRAS mutations as criteria for treatment with tyrosine kinase inhibitors: retro- and prospective observations in non-small-cell lung cancer. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO, 18*(1), 99-103.

Wakelee, H. A., Chang, E. T., Gomez, S. L., Keegan, T. H., Feskanich, D., Clarke, C. A., . . . West, D. W. (2007). Lung cancer incidence in never smokers. [Article]. *Journal of Clinical Oncology, 25*(5), 472-478. doi: 10.1200/jco.2006.07.2983

Weber, F., Fukino, K., Villalona-Calero, M., & Eng, C. (2005). Limitations of single-strand conformation polymorphism analysis as a high-throughput method for the detection of EGFR mutations in the clinical setting. [Letter]. *Journal of Clinical Oncology, 23*(24), 5847-5848. doi: 10.1200/jco.2005.01.5222

Weir, B. A., Woo, M. S., Getz, G., Perner, S., Ding, L., Beroukhim, R., . . . Meyerson, M. (2007). Characterizing the cancer genome in lung adenocarcinoma. *Nature, 450*(7171), 893-U822. doi: 10.1038/nature06358

Works, C. R., & Gallucci, B. B. (1996). Biology of lung cancer. *Seminars in oncology nursing, 12*(4), 276-284.

Wu, J. Y., Wu, S. G., Yang, C. H., Chang, Y. L., Chang, Y. C., Hsu, Y. C., . . . Yang, P. C. (2010). Comparison of gefitinib and erlotinib in advanced NSCLC and the effect of EGFR mutations. [Article]. *Lung Cancer, 72*(2), 205-212. doi: 10.1016/j.lungcan.2010.08.013

Xie, H.-G., & Frueh, F. W. (2005). Pharmacogenomics steps toward personalized medicine. *Personalized Medicine, 2*(4), 325-337.

Yano, T., Haro, A., Shikada, Y., Maruyama, R., & Maehara, Y. (2011). Non-small cell lung cancer in never smokers as a representative 'non-smoking-associated lung cancer': epidemiology and clinical features. [Review]. *International Journal of Clinical Oncology, 16*(4), 287-293. doi: 10.1007/s10147-010-0160-8

Yatabe, Y., & Mitsudomi, T. (2007). Epidermal growth factor receptor mutations in lung cancers. *Pathology international, 57*(5), 233-244.

Yoshida, T., Zhang, G. L., & Haura, E. B. (2010). Targeting epidermal growth factor receptor: Central signaling kinase in lung cancer. [Review]. *Biochemical Pharmacology, 80*(5), 613-623. doi: 10.1016/j.bcp.2010.05.014

Yu, S. R., Xie, L., Hou, Z. B., Qian, X. P., Yu, L. X., Wei, J., . . . Liu, B. R. (2011). Coamplification at lower denaturation temperature polymerase chain reaction enables selective identification of K-Ras mutations in formalin-fixed, paraffin-embedded tumor tissues without tumor-cell enrichment. [Article]. *Human Pathology, 42*(9), 1312-1318. doi: 10.1016/j.humpath.2010.06.018

Yun, C. H., Mengwasser, K. E., Toms, A. V., Woo, M. S., Greulich, H., Wong, K. K., . . . Eck, M. J. (2008). The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. [Article]. *Proceedings of the National Academy of Sciences of the United States of America, 105*(6), 2070-2075. doi: 10.1073/pnas.0709662105

Zhang, Z., Stiegler, A. L., Boggon, T. J., Kobayashi, S., & Halmos, B. (2010). EGFR-mutated lung cancer: a paradigm of molecular oncology. *Oncotarget, 1*(7), 497-514.

Zuo, Z., Chen, S. S., Chandra, P. K., Galbincea, J. M., Soape, M., Doan, S., . . . Luthra, R. (2009). Application of COLD-PCR for improved detection of KRAS mutations in clinical samples. *Modern Pathology, 22*(8), 1023-1031. doi: 10.1038/modpathol.2009.59

Appendix I

This table represents the known EGFR mutations, most of which are present at such a low frequency that their clinical significance is yet to be established.

Table 20 EGFR mutations in NSCLC. *Clinically significant mutations

Exon	Mutation type	Nucleotide change	Amino acid change
18	missense	c.2126A>C	p.E709A
		c.2155G>A	p.G719S
		c.2156G>C	p.G719A
		c.2170G>A	p.G724S
	del/insertion	c.2127_2130del4insC	p.E709_T710delinsD
19	del/insertion	c.2233_2245del15	p.K745_D749del
		c.2235_2249del15*	p.E746_A750del*
		c.2235_2249del15insTTC	p.E746_A750delinsF
		c.2236_2250del15	p.E746_A750del
		c.2237_2251del15	p.E746_A750del
		c.2237_22512del16insT	p.E746_A750delinsV
		c.2237_2238ins18	p.E746VinsPVAIKE
		c.2239_2248del10insC	p.L747_D749delinsP
		c.2239_2251del13insC	p.L747_T751delinsP
		c.2239_2258del20insCA	p.L747_P753delinsQ
		c.2240_2254del15	p.L747_T751del
		c.2240_2257del18	p.L747_P753delinsS
		c.2252_2276del25insA	p.T751_I759delinsN
20	missense	c.2303G>T	p.S768I
	del/insertion	c.2300_2308del9	p.A767_V769del
		c.2309_2310insCCAGCGTGG	p.D770_H773insGSVD
		c.2311A>G, 2312_2313insGGT	p.N771_P772insGY
		c.2317delCinsTACAACCCCT	p.H773_R776insYNPY
		c.2322_2323insCCACGT	p.C775_R776insPA
	silent mutation	c.2289C>T	p.A763A
		c.2313C>T	p.N771N
21	missense	c.2506C>T	p.R836C
		c.2573T>G*	p.L858R*

Appendix II

Table 21 List of trial abbreviations, definitions and description with trial outcome

Trial Abbreviation	Trial Definition	Description	Outcome	Reference
IDEAL	Iressa Dose Evaluation in Advanced Lung Cancer	Phase II clinical trial 210 patients with advanced NSCLC randomised to receive different doses of gefitinib	Similar efficacy in response rates, PFS and OS in both groups	(Ansari, et al., 2009)
INSTANA	IRESSA as Second-line Therapy in Advanced NSCLC - KoreA	Phase III 1217 Asian non-smoking patients with advanced NSCLC randomised to receive gefitinib or carboplatin and paclitaxel	Superior PFS for gefitinib compared with combination of carboplatin and paclitaxel	(C.-C. Lin & Yang, 2011)
INTACT	Iressa NSCLC Trials Assessing Combination Therapy	Phase III trial involving 1093 patients who received gefitinib and chemotherapy or chemotherapy alone.	No increase in survival benefit or improvement in response for combination therapy	(Ansari, et al., 2009; C.-C. Lin & Yang, 2011)
INTEREST	Iressa NSCLC Trials Evaluating Response and Survival versus Taxotere	Phase III trial with 1433 patients who had already received chemotherapy randomised to receive gefitinib or docetaxel	Increased overall survival in those patients receiving gefitinib but no significant difference in progression free survival. Sub-analysis of EGFR mutation positive patients showed increased PFS and response rates with gefitinib vs docetaxel.	(Cataldo, et al., 2011; C.-C. Lin & Yang, 2011)
IPASS	Iressa Pan-Asia Study	Randomised Phase III trial in Asia with non-smokers who received either gefitinib or chemotherapy with carboplatin and paclitaxel.	PFS at one year was significantly greater in those patients who received gefitinib (24.9% vs. 6.7%). OS similar in both groups but improvement in the quality of life in those receiving gefitinib.	(Cataldo, et al., 2011; Ma, Hui, & Mok, 2010; Mitsudomi, 2010)
ISEL	Iressa Survival Evaluation in	1692 patients with advanced NSCLC and prior	Compared with placebo, gefitinib had	Hann et al, (Masago, et

	advanced Lung cancer	chemotherapy therapy were randomised to receive gefitinib or placebo.	higher response rates (8.0% vs. 1.3%) but no change in survival benefit. Sub-analysis of non-smokers treated with gefitinib showed increased survival (8.9months vs. 6.1months with placebo.	al., 2008) (Wu, et al., 2010) (C.-C. Lin & Yang, 2011; Provencio, et al., 2009)
TALENT	Tarceva Lung Cancer Investigation Trial	Phase III involving 1172 patients who received chemotherapy with erlotinib or chemotherapy alone	Adding erlotinib to chemotherapy did not increase overall survival	(C.-C. Lin & Yang, 2011)
TRIBUTE	Tarceva Responses in Conjunction with Paclitaxel and Carboplatin	Randomised allocation of chemotherapy, placebo and erlotinib in 116 patients	After chemotherapy no significant differences were found between smokers and non-smokers in response or overall survival. The sub-group of non-smokers treated with erlotinib had a mean survival of 22.5months compared with only 10.1months for those allocated to placebo	(Provencio, et al., 2009)
BATTLE	Biomarker-Integrated Approaches of Targeted Therapy for Lung Cancer Elimination	255 pre-treated NSCLC patients were randomised to receive erlotinib, randetanib, or sorafenib based on the patients mutation status for *EGFR* and *KRAS*	48% disease control at 8 weeks (primary endpoint.	(E. S. Kim et al., 2011)
BR.21	National Cancer Institute of Canada Clinical Trial Group Trial	Randomised phase III trial of erlotinib vs. placebo involving 731 patients who had already received chemotherapy.	Erlotinib statistically significant improvement in OS and quality of life	(Cappuzzo, et al., 2005; Masago, et al., 2008; Provencio, et al., 2009)

PFS = progression free survival, OS = overall survival

Appendix III

The following alignments were a result of comparing the sequences obtained from the Waikato Sequencing Facility and the online database (http://blast.ncbi.nlm.nih.gov/Blast.cgi)

EGFR Exon 21 BLAST Alignments

NSCLC FFPE Specimen 34F3 Forward Sequence

```
Query  26   TCGCTTGGTGCACCGCGACCTGGCAGCCAGGAACGTACTGGTGAAAACACCGCAGCATGT  85
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  24   TCGCTTGGTGCACCGCGACCTGGCAGCCAGGAACGTACTGGTGAAAACACCGCAGCATGT  83

Query  86   CAAGATCACAGATTTTGGGCTGGCCAAACTGCTGGGTGCGGAAGAGAAAGAATACCATGC  145
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  84   CAAGATCACAGATTTTGGGCTGGCCAAACTGCTGGGTGCGGAAGAGAAAGAATACCATGC  143

Query  146  AGAAGGAGGCAAA  158
            |||||||||||||
Sbjct  144  AGAAGGAGGCAAA  156
```

NSCLC FFPE Specimen 34F3 Reverse Complement Sequence

```
Query  77   GGCATGAACTACTTGGAGGACCGTCGCTTGGTGCACCGCGACCTGGCAGCCAGGAACGTA  136
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    GGCATGAACTACTTGGAGGACCGTCGCTTGGTGCACCGCGACCTGGCAGCCAGGAACGTA  60

Query  137  CTGGTGAAAACACCGCAGCATGTCAAGATCACAGATTTTGGGCTGGCCAAAGTCTCTGGG  196
            |||||||||||||||||||||||||||||||||||||||||||||||||| |  |||||
Sbjct  61   CTGGTGAAAACACCGCAGCATGTCAAGATCACAGATTTTGGGCTGGCCAAACTG-CTGGG  119

Query  197  TGTGGAAGAGAACAGA  212
            || ||||||||| |||
Sbjct  120  TGCGGAAGAGAA-AGA  134
```

EGFR Exon 19 BLAST Alignments

NSCLC FFPE Specimen 34F3 Forward Sequence

```
Query  1    ATCTCACAATTGCCAGTTAACGTCTTCCTTCTCTCTCTGTCATAGGGACTCTGGATCCCA  60
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  634  ATCTCACAATTGCCAGTTAACGTCTTCCTTCTCTCTCTGTCATAGG-ACTCTGGATCCCA  692

Query  61   GAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  693  GAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCG  752

Query  121  AAAGCCAACAAGGAAATCCTCGAT  144
            |||||||||||||||||||||||
Sbjct  753  AAAGCCAACAAGGAAATCCTCGAT  776
```

NSCLC FFPE Specimen 34F3 Reverse Complement Sequence

```
Query  96   GGACTCTGGATCCCAGAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAAGGAATTAAGA  155
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    GGACTCTGGATCCCAGAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAAGGAATTAAGA  60

Query  156  GAAGCAACATCTCCGAAAGCCAACAAGGAAATCCTCGATGTG  197
            |||||||||||||||||||||||||||||||||||||||||
Sbjct  61   GAAGCAACATCTCCGAAAGCCAACAAGGAAATCCTCGATGTG  102
```

KRAS Exon 2 BLAST Alignments

NSCLC FFPE Specimen 34F3 Forward Sequence

```
Query  6    ATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGGCAAGAGTGCCTTGACGATACAGCT  65
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  190  ATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGGCAAGAGTGCCTTGACGATACAGCT  249

Query  66   AATTCAGAATCATTTTGTGGACGAATATGATCCAACAATAGAGG  109
            |||||||||||||||||||||||||||||||||||||||||||
Sbjct  250  AATTCAGAATCATTTTGTGGACGAATATGATCCAACAATAGAGG  293
```

NSCLC FFPE Specimen 34F3 Forward Sequence

```
Query  17   CCTCTATTGTTGGATCATATTCGTCCACAAAATGATTCTGAATTAGCTGTATCGTCAAGG  76
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  293  CCTCTATTGTTGGATCATATTCGTCCACAAAATGATTCTGAATTAGCTGTATCGTCAAGG  234

Query  77   CACTCTTGCCTACGCCACCAGCTCCAACTACCACAAGTTTATATTCAGTCATTTTCAGCA  136
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  233  CACTCTTGCCTACGCCACCAGCTCCAACTACCACAAGTTTATATTCAGTCATTTTCAGCA  174

Query  137  GGCCT  141
            |||||
Sbjct  173  GGCCT  169
```

Appendix III

Table 22 Sequencing Data from Standard PCR and direct sequencing. Mutation hotpots in respective exons are in red (KELREA deletion in exon 19, L858R in exon 21 and codon 12 and 13 SNPs in KRAS)

Exon 19 Deletions																			
EGFR protein	**739**	**K**	**I**	**P**	**V**	**A**	**I**	**K**	**E**	**L**	**R**	**E**	**A**	**T**	**S**	**P**	**K**	**A**	**N**	**756**
EGFR gene	**2215**	**AAA**	**ATT**	**CCC**	**GTC**	**GCT**	**ATC**	**AAG**	**GAA**	**TTA**	**AGA**	**GAA**	**GCA**	**ACA**	**TCT**	**CCG**	**AAA**	**GCC**	**AAC**	**2268**
24A6		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
54A4		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
98F7		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
77B4		AAA	ATT	CCC	GTC	GCT	ATC	AAG	_____	____	_____	_____	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
34F3		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
19A8		___	___	___	___	___	___	___	___	___	___	___	___	___	___	___	___	___	___	
Exon 21 Mutations																				
EGFR Protein	**850**	**H**	**V**	**K**	**I**	**T**	**D**	**F**	**G**	**L**	**A**	**K**	**L**	**L**	**G**	**853**				
EGFR gene	**2538**	**CAT**	**GTC**	**AAG**	**ATC**	**ACA**	**GAT**	**TTT**	**GGG**	**CTG**	**GCC**	**AAA**	**CTG**	**CTG**	**GGT**	**2589**				
24A6		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT					
54A4		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT					
98F7		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT					
77B4		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT					
34F3		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT					
19A8		___	___	___	___	___	___	___	___	___	___	___	___	___	___					
KRAS Exon 2 Mutations																				
KRAS protein	**1**	**M**	**T**	**E**	**Y**	**K**	**L**	**V**	**V**	**V**	**G**	**A**	**G**	**G**	**V**	**G**	**K**	**S**	**A**	**18**
KRAS gene	**1**	**ATG**	**ACT**	**GAA**	**TAT**	**AAA**	**CTT**	**GTG**	**GTA**	**GTT**	**GGA**	**GCT**	**GGT**	**GGC**	**GTA**	**GGC**	**AAG**	**AGT**	**GCC**	**54**

24A6	ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC
54A4	ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC
98F7	ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC
77B4	ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC
34F3	ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC
19A8	ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC

Table 23 Sequencing Data following COLD-PCR and direct sequencing

Exon 19 Deletions																			
EGFR protein	**739**	**K**	**I**	**P**	**V**	**A**	**I**	**K**	**E**	**L**	**R**	**E**	**A**	**T**	**S**	**P**	**K**	**A**	**N**	**756**
EGFR gene	**2215**	**AAA**	**ATT**	**CCC**	**GTC**	**GCT**	**ATC**	**AAG**	**GAA**	**TTA**	**AGA**	**GAA**	**GCA**	**ACA**	**TCT**	**CCG**	**AAA**	**GCC**	**AAC**	**2268**
24A6		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
54A4		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
98F7		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
77B4		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
34F3		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
19A8		AAA	ATT	CCC	GTC	GCT	ATC	AAG	GAA	TTA	AGA	GAA	GCA	ACA	TCT	CCG	AAA	GCC	AAC	
Exon 21 Mutations																			
EGFR Protein	**850**	**H**	**V**	**K**	**I**	**T**	**D**	**F**	**G**	**L**	**A**	**K**	**L**	**L**	**G**	**853**			
EGFR gene	**2538**	**CAT**	**GTC**	**AAG**	**ATC**	**ACA**	**GAT**	**TTT**	**GGG**	**CTG**	**GCC**	**AAA**	**CTG**	**CTG**	**GGT**	**2589**			
24A6		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT				
54A4		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT				
98F7		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT				
77B4		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT				

34F3		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT				
19A8		CAT	GTC	AAG	ATC	ACA	GAT	TTT	GGG	CTG	GCC	AAA	CTG	CTG	GGT				
KRAS Exon 2 Mutations																			
KRAS protein	**1**	**M**	**T**	**E**	**Y**	**K**	**L**	**V**	**V**	**V**	**G**	**A**	**G**	**G**	**V**	**G**	**K**	**S**	**A**	**18**
KRAS gene	**1**	**ATG**	**ACT**	**GAA**	**TAT**	**AAA**	**CTT**	**GTG**	**GTA**	**GTT**	**GGA**	**GCT**	**GGT**	**GGC**	**GTA**	**GGC**	**AAG**	**AGT**	**GCC**	**54**
24A6		ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC	
54A4		ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC	
98F7		ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC	
77B4		ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC	
34F3		ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC	
19A8		ATG	ACT	GAA	TAT	AAA	CTT	GTG	GTA	GTT	GGA	GCT	GGT	GGC	GTA	GGC	AAG	AGT	GCC	